Nos. 15-1080, -1081, -1082, -1083, -1084, -1085, -1086, -1087, -1088, -1089, -1090, -1092, -1093, -1094, -1095, -1096, -1097, -1098, -1099, -1100, -1101

IN THE

United States Court of Appeals

FOR THE FEDERAL CIRCUIT

MCRO, INC., DBA PLANET BLUE,

Plaintiff-Appellant,

v.

BANDAI NAMCO GAMES AMERICA INC., NAUGHTY DOG, INC., KONAMI DIGITAL ENTERTAINMENT, INC., SEGA OF AMERICA, INC., ELECTRONIC ARTS INC., OBSIDIAN ENTERTAINMENT, INC., DISNEY INTERACTIVE STUDIOS, INC., SQUARE ENIX, INC., NEVERSOFT ENTERTAINMENT, INC., TREYARCH CORPORATION, CAPCOM USA, INC., SONY COMPUTER ENTERTAINMENT AMERICA LLC, ATLUS U.S.A., INC., SUCKER PUNCH PRODUCTIONS, LLC, INFINITY WARD, INC., LUCASARTS, A DIVISION OF LUCASFILM ENTERTAINMENT COMPANY LTD. LLC, WARNER BROS. INTERACTIVE ENTERTAINMENT, A DIVISION OF WARNER BROS. HOME ENTERTAINMENT INC., ACTIVISION PUBLISHING, INC., AND BLIZZARD ENTERTAINMENT, INC.,

Defendants-Appellees,

and

VALVE CORPORATION, CODEMASTERS USA GROUP, INC., CODEMASTERS SOFTWARE INC., CODEMASTERS, INC., AND THE CODEMASTERS SOFTWARE COMPANY LIMITED,

Defendants-Appellees.

Appeal from the United States District Court
for the Central District of California

CORRECTED BRIEF FOR PLAINTIFF-APPELLANT MCRO, INC.

John Whealan
4613 Merivale Road
Chevy Chase, MD 20815
(202) 994-2195 (telephone)

Jeffrey A. Lamken
Counsel of Record
Michael G. Pattillo, Jr.
MOLO LAMKEN LLP
The Watergate, Suite 660
600 New Hampshire Avenue, N.W.
Washington, D.C. 20037
(202) 556-2000 (telephone)
(202) 556-2001 (fax)
jlamken@mololamken.com

Counsel for Plaintiff-Appellant McRO, Inc.
(Additional counsel listed on inside cover)

Mark S. Raskin
Robert A. Whitman
John F. Petrsoric
MISHCON DE REYA NEW YORK LLP
750 Seventh Avenue 26th Floor
New York, NY 10019
(212) 612-3279 (telephone)
(212) 612-3297 (fax)
mark.raskin@mishcon.com

Counsel for Plaintiff-Appellant
McRO, Inc.

Form 9

FORM 9. Certificate of Interest

UNITED STATES COURT OF APPEALS FOR THE FEDERAL CIRCUIT

McRO, Inc., DBA Planet Blue v. Bandai Namco Games America Inc.

No. 15-1080; 15-1099

CERTIFICATE OF INTEREST

Counsel for the (petitioner) (appellant) (respondent) (appellee) (amicus) (name of party) McRO, Inc., DBA Planet Blue certifies the following (use "None" if applicable; use extra sheets if necessary):

1. The full name of every party or amicus represented by me is:

McRO, Inc., DBA Planet Blue

2. The name of the real party in interest (if the party named in the caption is not the real party in interest) represented by me is:

N/A

3. All parent corporations and any publicly held companies that own 10 percent or more of the stock of the party or amicus curiae represented by me are:

N/A

4. ☑ The names of all law firms and the partners or associates that appeared for the party or amicus now represented by me in the trial court or agency or are expected to appear in this court are:

Mark S. Raskin, Robert Whitman, John Petrsoric & Eric Berger of Mishcon de Reya New York LLP; Jeffrey A. Lamken & Michael G. Pattillo of Molo Lamken LLP; John Whealan

February 27, 2015
Date

/s/ Jeffrey A. Lamken
Signature of counsel

Jeffrey A. Lamken
Printed name of counsel

Please Note: All questions must be answered
cc: Counsel of Record

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

Cases

Statutes

Other Authorities

STATEMENT OF RELATED CASES PER FEDERAL CIRCUIT RULE 47.5

Five related cases are stayed in the U.S. District Court for the District of Delaware pending this appeal: *McRO Inc. v. Bethesda Softworks, LLC*, No. 12-cv-01509; *McRO Inc. v. Harmonix Music Sys., Inc.*, No. 12-cv-01510; *McRO Inc. v. Rockstar Games, Inc.*, No. 12-cv-01513; *McRO Inc. v. Take-Two Interactive Software, Inc.*, No. 12-cv-01517; and *McRO Inc. v. 2K Games, Inc.*, No. 12-cv-01519.

JURISDICTIONAL STATEMENT

This appeal arises from a decision of the U.S. District Court for the Central District of California. The district court had jurisdiction under 28 U.S.C. §§ 1331 and 1338(a). The district court granted defendants judgment on the pleadings on September 22, 2014. Plaintiff timely filed a notice of appeal on October 22, and the district court entered final judgment on October 31. This Court has jurisdiction under 28 U.S.C. § 1295.

ISSUE PRESENTED

Whether the district court erred in holding that the claimed "method for automatically animating lip synchronization and facial expression of three-dimensional characters," A37, 11:27-28, covers "abstract ideas" that are not patent-eligible under 35 U.S.C. § 101.

STATEMENT OF THE CASE

In 2012 and 2013, McRO, Inc., d/b/a Planet Blue ("Planet Blue"), commenced the instant actions under 35 U.S.C. § 271, alleging that defendants infringe

U.S. Patent Nos. 6,307,576 ("the '576 Patent") and 6,611,278 ("the '278 Patent"). The actions were ultimately consolidated in the U.S. District Court for the Central District of California. One set of cases was consolidated under Case No. 12-cv-10322, and another set under Case No. 13-cv-1874.

On July 10, 2014, all defendants moved for judgment on the pleadings, urging that the patents claim patent-ineligible subject matter under 35 U.S.C. § 101. On September 22, the district court granted the motion.

STATEMENT OF FACTS

Maury Rosenfeld is an award-winning animator and the founder of plaintiff-appellant Planet Blue, a visual effects company. The patents-in-suit are directed to the field of 3-D computer animation. They teach methods, embodied in computer software, for automatically manipulating the facial features of a computer-generated character to make a video of the character realistically speaking pre-recorded dialogue. Rosenfeld's patented method has been hailed as "revolutionary." Before these suits, many of the named defendants hired Planet Blue to do animation projects for them.

Defendants (video-game makers) began using the patented techniques without a license. In turn, Planet Blue sued for infringement. After claim construction and motions to dismiss, defendants moved for judgment on the pleadings, arguing

that the patents cover "abstract ideas" that are not patent-eligible under 35 U.S.C. § 101.

The district court found that the claims "do ***not*** seem directed to an abstract idea." A13 (emphasis added). To the contrary, "[t]hey are ***tangible***, each covering an approach to ***automated three-dimensional computer animation***, which is a ***specific technological process***." *Id.* (emphasis added). However, the court believed it had to divide the claims up for § 101 analysis so that it could focus solely on the invention's "point of novelty." A14-17; *but see Diamond v. Diehr*, 450 U.S. 175, 188-90 (1981). The court then created a claim chart (as in a § 103 obviousness analysis), identifying every step that existed in the prior art—*i.e.*, computer-animation techniques that the method was automating—and stripped those steps from the claims. A17-18. According to the court, that left only "the idea of using rules, including timing rules, to automate the process of generating keyframes." A17. The court held that that was an unpatentable abstract idea. A18-19.

I. The Field of the Invention: Computer-Generated 3-D Animation

The patents-in-suit are addressed to the field of 3-D computer animation: They teach a rapid, cost-effective method for providing automated lip synchronization for animated characters. The method starts with a specialized transcript of pre-recorded dialogue and a model of the character's face. It then automatically moves the model's facial features, creating a series of images—essentially, a

video—that realistically depicts the model speaking in sync with the dialogue. The patented method overcomes significant deficiencies of prior-art lip-synchronization methods, which were either time-consuming and expensive, or produced low-quality, unrealistic animation. Defendant Warner Brothers called Planet Blue's technique "revolutionary," explaining that the "process offers high-quality lip sync with a fast turnaround, compatible with a range of 3-D production software." A4995.

Planet Blue submitted a video tutorial for the district court that covers background principles of computer-generated animation and the invention. The video is included on a CD-ROM in the appendix at A3573. We encourage the Court to view the video. Critical excerpts are also provided on another CD-ROM in the appendix at A6121.[1] Where animation is concerned, a written account is a poor substitute for seeing the real thing.

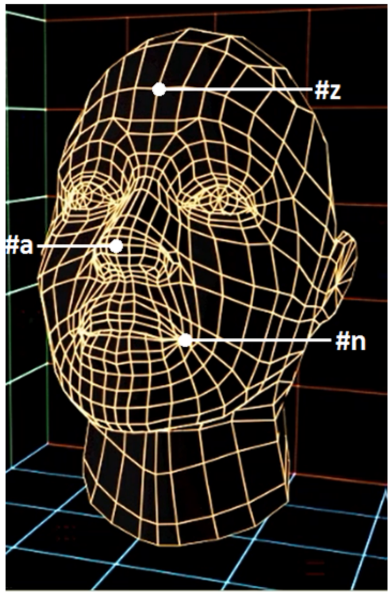
A. Background Principles of Computer-Generated Facial Animation

1. *The Neutral Model*

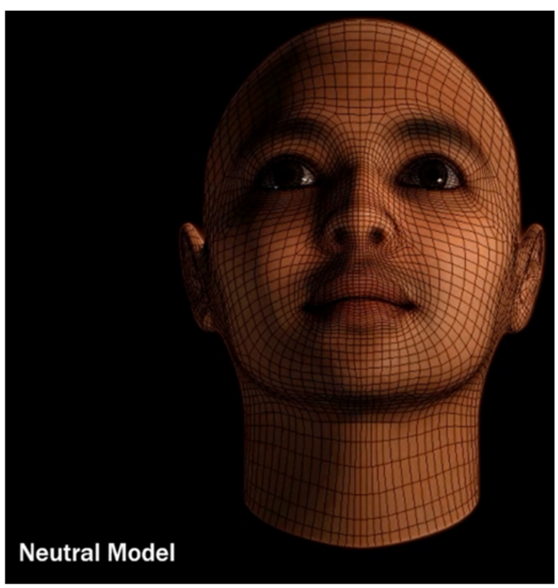
Computer-generated animation begins with an artist creating a character. The starting point is the "neutral model"—a reference model used to depict the character's face with a neutral expression. A32, 1:50-51; A35, 7:52-53; A3573, 1:20. One foundation of the neutral model is the "mesh" (shown below).

---

[1] The tutorial is cited as "A3573, [minute]:[second]." The excerpts are cited as "A6121, Clip __."



A3573, 1:25 (annotations added). The mesh is defined by "vertices"—points in three-dimensional space that outline the model's surface. A3573, :52. In the image above, vertex #n corresponds with the left corner of the mouth, vertex #a with the tip of the nose, and vertex #z with the center of the forehead. After the mesh is created, the animator adds details like skin, lips, and shading. Computers can easily manipulate models based on meshes.



A3573, 1:39, 7:19.

2. *Phonemes and Visemes*

The patents are directed to a method for depicting speech in 3-D computer animation. Two terms concerning speech are critical to understanding that process: “phonemes” and “visemes.”

Phonemes: Phonemes are the “smallest units of speech.” A32, 1:35. A phoneme “corresponds to a single sound,” such as “th,” “aah,” “ee,” or “oh.” A32, 1:35-36; A3573, 2:56. There are standardized “alphabets” of phonemes. A32, 1:36-39.

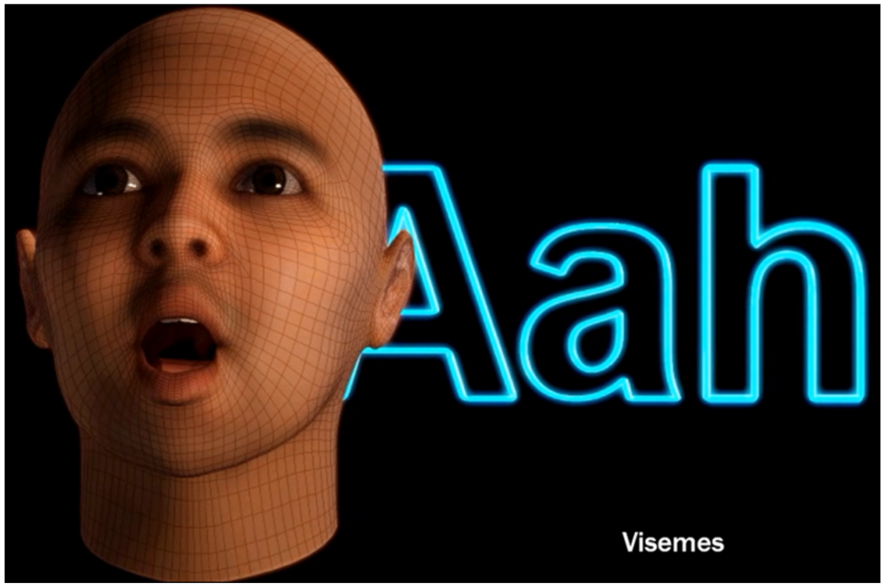
The goal of lip-synchronization animation is to match the character’s facial expression to spoken dialogue. Some processes, such as the method here, utilize a transcript of that dialogue (other techniques, like facial capture, do not). Instead of listing words, however, the transcript lists the phonemes that are spoken. It also indicates the point in time that each phoneme occurs. This is called a “time aligned phonetic transcription” or “TAPT.” A32, 1:32-34. The TAPT (illustrated below) can be generated by automatic speech-recognition programs. A32, 1:43-44.

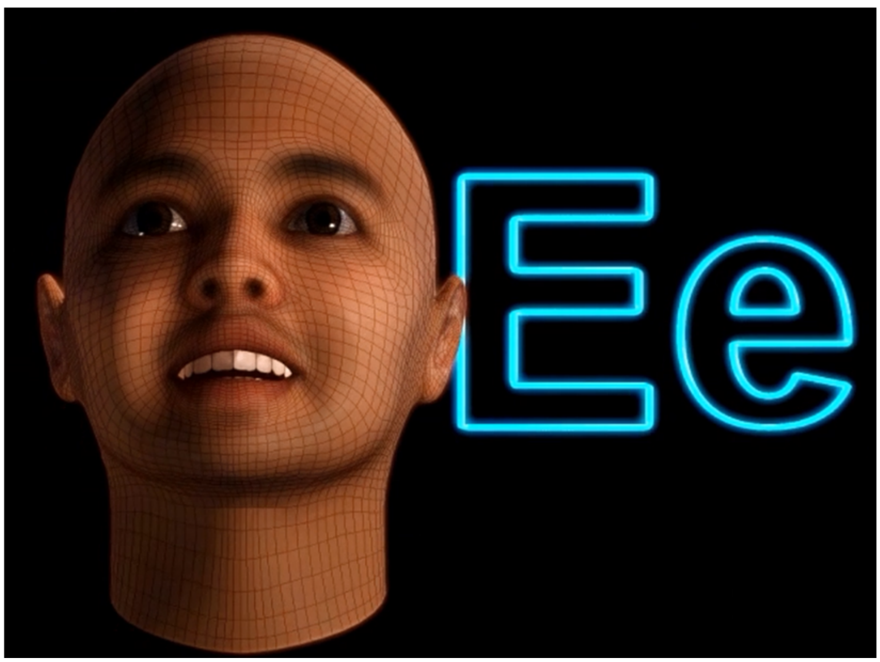
time (sec)	phoneme	word
0		Sil
1.895	h	hello
1.965	eh	
1.995	l	
2.105	o	
2.137	w	
2.165	dh	there
2.235	eh	
2.335	r	
2.435	sil	
2.475	h	how
2.545	a	
2.601	w	
2.635	AA	are
2.66	r	
2.695	y	you
2.835	uw	
2.885	t	today
2.945	ah	
2.985	d	
3.045	e	
3.16	y	
3.225	sil	

A3573, 4:14.

Visemes: A viseme is an image of a face speaking a certain phoneme. A3573, 4:24. Because the face looks the same when making certain sounds, such as "F" and "V," the same viseme can sometimes be used for multiple phonemes. A3573, 4:50.

In computer animation, the artist starts with the neutral model and then creates a library of visemes showing what that character looks like when speaking the various sounds. Below are visemes showing the model on p. 5, *supra*, speaking the "Aah" and "Ee" phonemes:





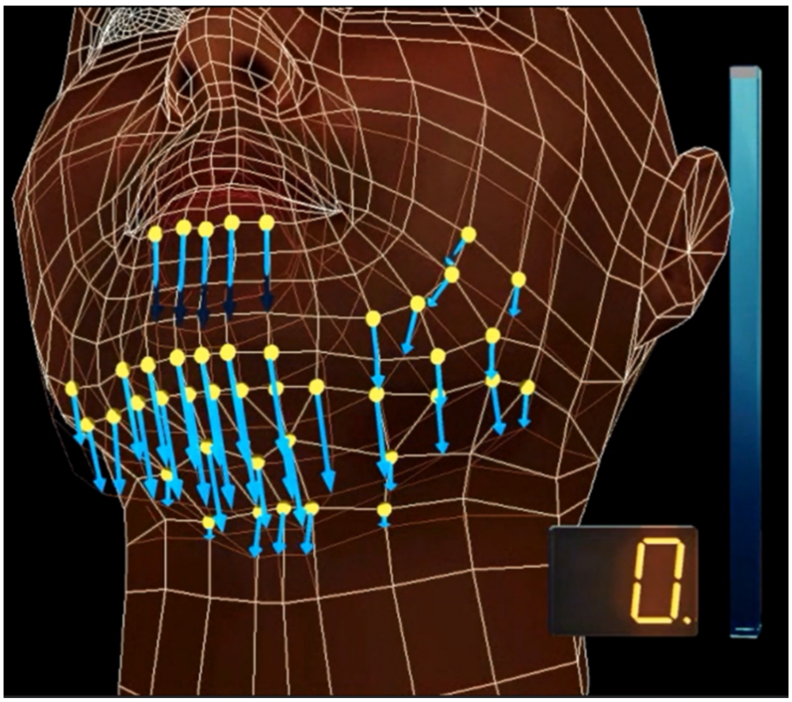
A3573, 7:24, 7:31.

3. *Transforming the Neutral Model Into "Morph Targets"*

In computer animation, visemes are sometimes called "morph targets." A morph target is a model representing a specific state—in the case of a viseme, a

facial expression—that the animator will transform, or "morph," the neutral model into. A3573, 7:00, 7:31; A32, 1:49-51.

A morph target is created by manipulating the vertices of the neutral model to form a different facial shape. Accordingly, a morph target has the same number of vertices as the neutral model, and each vertex on the morph target corresponds to a vertex on the neutral model. A32, 1:51-53. For example, a vertex labeled #n may represent the left corner of the mouth on the neutral model and on each morph target, but its position may vary depending on the facial expression depicted. A32, 1:53-56. Thus, while each model is a different shape, they share an identical structure. The following diagram illustrates how vertices on the neutral model can be moved to a different position to create the effect of the mouth opening:



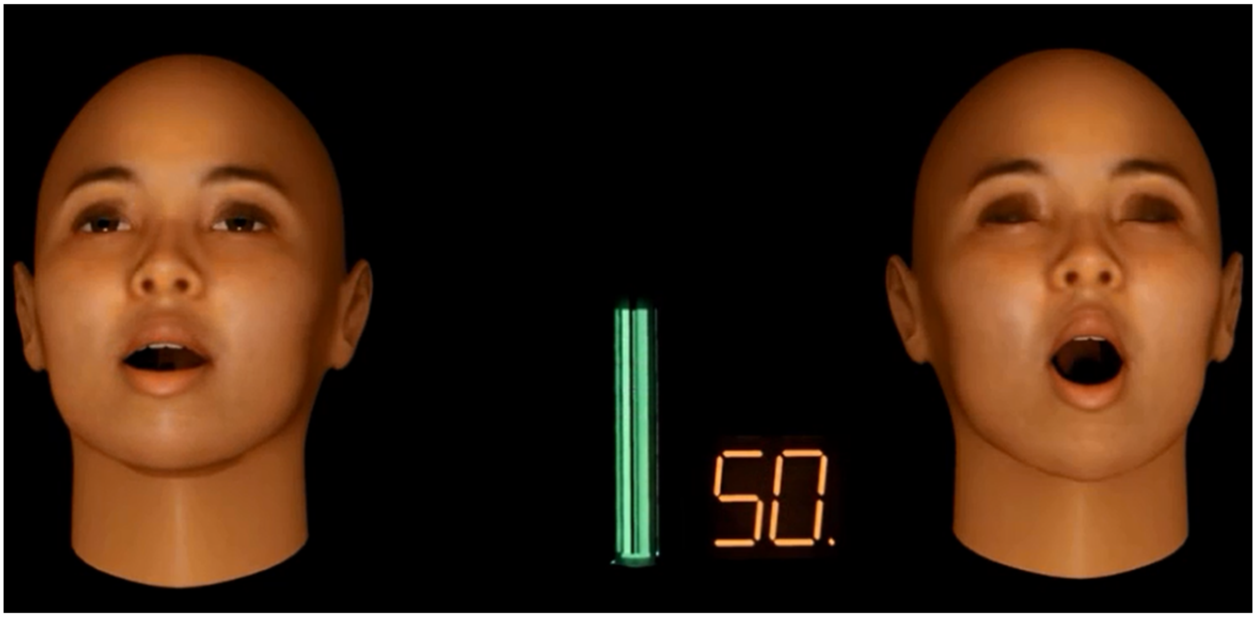
A3573, 15:44.

A morph target can be represented as a "delta set" (or a set of "delta vectors"). A32, 1:58-62; A3573, 14:10. The "deltas" are the differences in position of each vertex on the morph target from the neutral model (represented by the arrows in the diagram above). A32, 1:58-62.

Once a morph target has been defined, an animator can use "morph weights" to transform the neutral model only part-way toward that morph target. A32, 1:63-2:15. A morph weight is a value, between 0 (0%) and 1 (100%), that is applied to the delta set to determine how far the neutral model's vertices should be moved toward the morph target. A32, 1:63-65; A3573, 6:08, 7:56, 16:16, 21:14.

A morph weight represents the degree to which the facial shape corresponding to the morph target is expressed. For example, if the morph target for the "oh" viseme is assigned a morph weight of 1, the neutral model would be modified to look exactly like the "oh" viseme. A32, 2:16-19. But if the "oh" viseme is assigned a morph weight of .5, the neutral model's vertices would be moved only halfway toward the "oh" viseme. A32, 2:16-22.

In the illustration below, the model on the right shows the "oh" viseme; the left is the "oh" viseme with a .5 morph weight. The mouth on the left is open halfway compared with the mouth on the right.



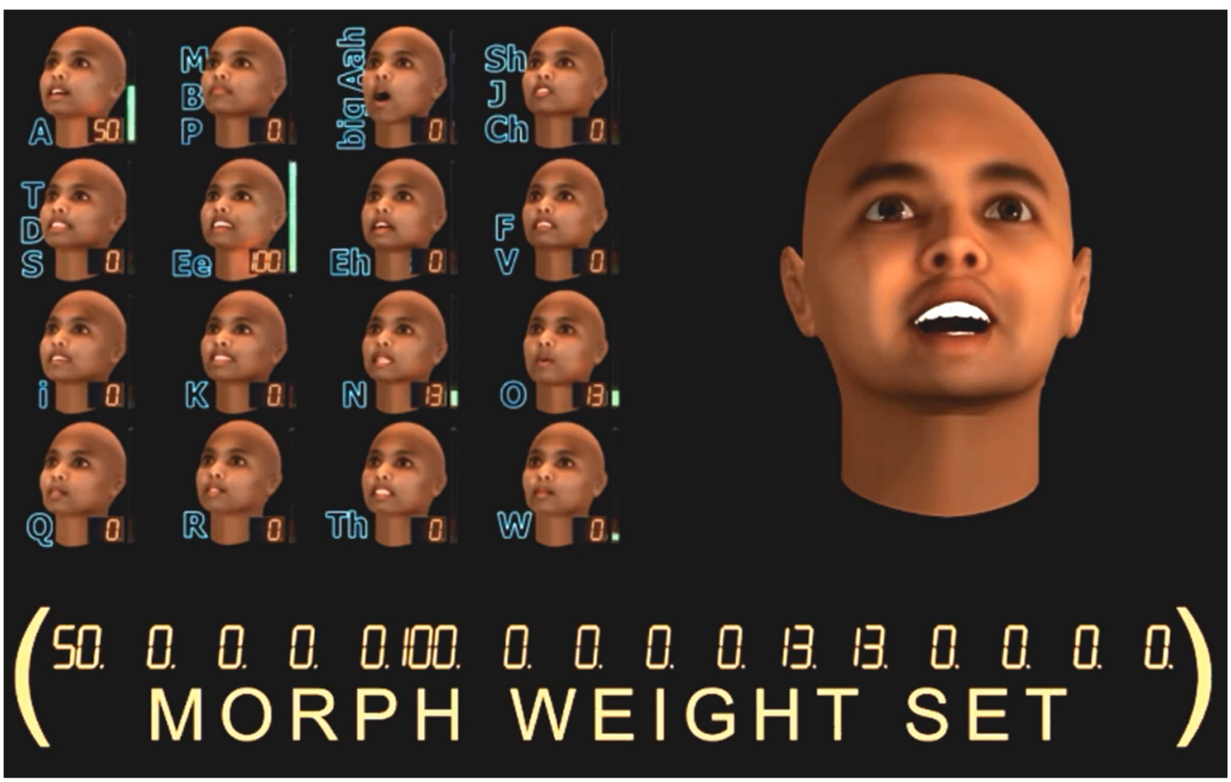
A3573, 7:49.

Several morph targets can be blended to create a model that exhibits characteristics of each. A32, 2:23-28. For example, an animator could simultaneously assign the "oh" viseme a .3 morph weight and the "ee" viseme a .7 morph weight. The result would be an expression that reflects some of the "oh" model's "openness," but more of the "ee" model's characteristic of being stretched horizontally. A32, 2:23-27.

How much of each morph target the animator chooses to "mix in" a blended model can be represented as a "morph weight set." A33, 4:39-41; A3573, 11:50-11:59. A morph weight set is a set of values in which each entry represents the

weight of one of the morph targets (visemes). A3573, 11:18, 16:35, 21:20; A33, 4:39-41. As the district court explained, when a morph weight set is "applied" to the neutral model, it "transform[s]" the model to a particular facial expression. D.Ct.Dkt.#298-1, at 9.

Consider an example where the animator has created a neutral model and sixteen visemes/morph targets. The animator may wish to depict a facial expression where the "Ee" viseme is strongly expressed with a 100% morph weight; the "A" viseme is significantly expressed with a 50% morph weight; others, such as "N" and "O," are more subtle with a morph weight of 13%; and others are not expressed at all. The following illustration shows the resulting facial expression and its morph weight set:



A3573, 11:48. Assigning a morph weight set to a specific time in the animation is called "setting a keyframe."

B. Prior-Art Methods of Lip Synchronization

There were two main prior-art methods of lip synchronization in 3-D computer-generated animation. Each had significant limitations.

1. *The Manual Approach*

The so-called "manual" approach begins with an artist creating a neutral model of the character and a library of visemes (*i.e.*, morph targets). A32, 1:44-51. The manual approach is not to be confused with hand-drawn animation—it uses a computer and animation software to create and manipulate the character models. A32, 1:44-47. What makes it "manual" is that a human artist sits at a computer and, relying on her own artistic judgment, manipulates the model to match the speech in the time-aligned transcript of the dialogue.

In this method, the artist creates a computer image of the model making a particular facial expression for each of the most important moments—these are the "keyframes." *See* A32, 2:29-34. For each keyframe, the artist chooses morph targets (visemes) from the library and adjusts "sliders" that control how much each morph target is emphasized until the character model on the screen reflects the desired blend of facial expressions. *See* A3573, 11:00 (demonstrating use of sliders); A4949 ¶9. The software derives a morph weight set that corresponds to

the sliders' positions. *See* p. 12, *supra* (illustration); A3573, 11:19-47; A6121, Clip 1.

Keyframes depict only the important moments in the animation. While the artist is guided by the time-aligned phonetic transcript, the artist decides where in time to set keyframes—she does not simply enter a keyframe at each phoneme. An animator must create an extremely large number of keyframes to accurately depict speech. And creating the illusion of smooth movement requires even more images—30 images per second for standard television. A35, 7:15-18. Traditional animation software provides those additional images by "interpolat[ing]" between keyframes. A32, 2:33-34.

The manual process yields high-quality animation, because a skilled artist decides when in time to set keyframes, and tweaks the morph weight sets at each keyframe to achieve precisely the desired effect. But the process is also laborious, time-consuming, and expensive. A32, 2:34-35. The manual process remains the method of choice for major motion pictures and other projects where quality is paramount and time and budget are not major constraints.

2. *Prior Attempts at Automation*

A second, more automated, method of lip synchronization also existed in the prior art. It, too, began with an artist creating a neutral model of the character and a library of visemes/morph targets. A3573, 8:24. But from there, software takes

over. It automatically sets a keyframe at each time a phoneme appears in the time-aligned transcript. A3573, 8:42-9:00. That keyframe, however, is simply the viseme that corresponds to the phoneme—morph targets are not blended to create a nuanced expression. *Id.* Because each viseme is given a morph weight of 1 or 0, with nothing in-between, the viseme is either fully expressed or not expressed at all. *Id.* The software then uniformly interpolates between those keyframes to generate an image at each frame of the animation. A3573, 9:02.

That automated process is much cheaper than manually setting keyframes, but it yields poor-quality animation. A3573, 9:33, 17:47. Mechanically applying visemes without blending gives the animation a "flappy" quality—the expression of the visemes appears artificial and over-articulated. A3573, 9:40-55, 17:51-18:12. And the faster the tempo of speech, the more pronounced the "flappy" effect. A3573, 9:40-55, 17:58. Those shortcomings are demonstrated in the video tutorial addressing the prior art. A3573, 9:40-55, 17:51-18:12; A6121, Clip 2. The quality of animation produced by this method is rarely acceptable even for time- and budget-sensitive projects, and often requires extensive after-the-fact "tuning" by an artist. *See* A3573, 9:58.

II. The Patents-in-Suit: Maury Rosenfeld's Automated Method for High-Quality 3-D Lip Synchronization and Facial Animation

A. The Inventor and His Insight

Maury Rosenfeld has worked as a successful computer graphics and visual effects designer for over 30 years. *See* D.Ct.Dkt.#73, ¶7. He has done special-effects work for shows such as *Star Trek: The Next Generation*, *Max Headroom*, and *Pee Wee's Playhouse*. *Id*. Rosenfeld is active in the field of computer imaging, and his work is regularly seen by millions. He has earned numerous awards, including an Emmy, a Monitor Award, and an award from the National Computer Graphics Association. *Id*. Rosenfeld founded Planet Blue, a visual effects company, in 1988. *Id.* ¶6.

An accomplished animator, software engineer, and small-business owner, Rosenfeld was familiar with the expense-vs.-quality tradeoffs in the prior-art methods of 3-D computer-generated animation. He sought to develop software that automates the lip-synchronization process to produce high-quality, realistic animation at less time and cost.

Rosenfeld realized that prior efforts to automate lip synchronization failed to account for the fact that speakers do not express each viseme to an equal degree, and do not transition between visemes evenly. A3573, 19:05. Rather, the extent to which any viseme is expressed is influenced by the context (surrounding visemes) and the pace (rate) of speech. A3573, 19:26. For example, a face might fully ex-

press the "l" viseme when it begins a word, as in "love." But the extent to which the "l" viseme is expressed decreases in relation to how quickly it follows certain "closed mouth" visemes, such as the "b" viseme. Consider a person speaking the words "blob" and "Bob." Because the "l" viseme comes right after the closed-mouth "b" viseme when saying "blob," a person speaking at a normal pace may hardly express the "l" viseme at all. But when the same words are said slowly and deliberately, they look very different. A3573, 19:20. In the manual approach, an artist will account for such things, manipulating the model until it looks "right." But the automated method ignored such nuances and thus produced unnatural results.

Rosenfeld's critical insight was that, if one could configure software to account for just those specific variables—the context and pace of speech—one could produce high-quality animation automatically. *See* A4950 ¶12; A4953 ¶20. Rosenfeld invented a groundbreaking method that achieves that goal. Unlike prior automated methods, his invention does not rotely depict the visemes that correspond to the spoken phonemes; it also automatically weights and blends the visemes based on the context and pace of the speech. The method is rapid and cost-effective, A32, 2:50-54, but yields sufficiently realistic animation for applications like cartoons and video games. A32, 2:49-50; A33, 3:22-23. The many videos he has created demonstrate this in practice. *See, e.g.*, A3573, :01-:25; A6121, Clip 3.

Rosenfeld's innovations represented a tremendous advance. An internal memorandum from defendant Warner Brothers called Planet Blue's lip-sync technique "revolutionary," stating that the "process offers high-quality lip sync with a fast turnaround, compatible with a range of 3-D production software." A4995. And defendants Disney, LucasArts, Electronic Arts, and Warner Brothers hired Planet Blue to work on animation and lip-synchronization projects after seeing demonstrations and discussing the technology with Rosenfeld. *See, e.g.*, A4978-93.

B. The Patented Technology

The '576 Patent, titled "Method for Automatically Animating Lip Synchronization and Facial Expression of Animated Characters," issued to Rosenfeld on October 23, 2001. A27. The '278 Patent, a continuation of the '576 Patent, issued on August 26, 2003. A39. The patents teach a "method embodied in computer software" that automatically moves a model's facial features to make a video of the character speaking in sync with pre-recorded dialogue. A32, 2:40-44. It begins with a phonetic transcript of recorded dialogue and a 3-D model of a character's face, and applies rules that evaluate the transcript in terms of "phoneme sequence and time of said phoneme sequence." A37, 11:31-33. The result is "lip synchronization and facial expression control of . . . animated characters," such that

the computer-generated video is synchronized with the audio dialogue. A37, 11:45-47.

1. *The Claims*

Claim 1 of the '576 Patent is representative of the asserted claims. It recites:

> A method for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:
>
> obtaining a first set of rules that define output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence;
>
> obtaining a timed data file of phonemes having a plurality of subsequences;
>
> generating an intermediate stream of output morph weight sets and a plurality of transition parameters between two adjacent morph weight sets by evaluating said plurality of subsequences against said first set of rules;
>
> generating a final stream of output morph weight sets at a desired frame rate from said intermediate stream of output morph weight sets and said plurality of transition parameters; and
>
> applying said final stream of output morph weight sets to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters.

A37, 11:26-47. The patents also include apparatus claims encompassing a computer system programmed to perform the method. *See, e.g.*, A37, 12:29-54.

2. *The Operation of the Invention*

While the patented invention represents a leap forward, it of course builds on prior-art principles and techniques of 3-D computer-generated animation. For example, it requires a time-aligned phonetic transcript of the recorded dialogue.

See A32, 1:32-34. It also requires an artist-created neutral model of the character and library of morph targets/visemes. *See* A32, 1:47-51; A35, 7:37-40; A3573, 7:18, 7:35; pp. 4-8, *supra*.

What sets the invention apart is how it determines the facial movements that will be synchronized with the audio dialogue. Among other things, it uses rules that evaluate the TAPT and—depending on the sequence of phonemes, the context of each phoneme with respect to surrounding phonemes, and the pace of speech—***automatically*** determines when to set keyframes, the appropriate morph weight set at each keyframe, and how the animation should transition between keyframes.

a. *Rules Based on Timing and Context of the Speech*

As noted above, the invention allows for automated lip synchronization using specific rules based on timing and context. "Correspondence" rules establish the morph weight set to be applied—and thus, the facial expression the character should make—when a particular sub-sequence of phonemes is encountered in the time-aligned transcript of the dialogue. A33, 4:55-58. And "transition" rules set the timing of transitions from one morph weight set to the next—again depending on the particular phoneme sub-sequence encountered in the TAPT. A33, 4:58-61; A34, 6:51-59.

The patents put correspondence rules and transition rules into three categories: default rules, secondary rules, and post-processing rules. A34, 5:9-11.

Default correspondence rules are the morph weight sets for the visemes. A34, 6:46-51. An example would be that, if the software encounters an "h" as in "house" in the phonetic transcript, it should use the morph weight set corresponding to that viseme at the time the phoneme occurs in the TAPT. A35, 7:60-63. A default transition rule might create a visual hold on the phoneme for the majority of its duration, and allow transitions at the beginning and end of the phoneme times proportional to their durations. *See, e.g.*, A35, 8:9-14. The default correspondence and transition rules must be "complete enough to create valid output for any [phoneme sub-sequence] encountered at any point in the TAPT." A34, 5:10-13.

Secondary rules have more complex criteria. A secondary rule might be based on a phoneme's context (what phonemes precede and follow it), its duration, and/or the duration of adjacent phonemes. A34, 5:33-37. A secondary correspondence rule, for example, might account for the "blob/Bob" issue discussed above (at p. 17) by specifying a morph weight set that suppresses the "l" viseme when it quickly follows a closed-mouth viseme such as "b." *See* A35, 8:21-23 (discussing different examples of secondary correspondence rules based on phoneme sequence). And a secondary transition rule might account for how the mouth looks when speech stops, altering the transition time where a phoneme precedes a

silence. A35, 8:29-31. Where the criteria for a secondary rule are met, it displaces otherwise applicable default rules. A34, 5:23-26.

Post-processing rules further manipulate the morph weight sets after the default and secondary rules have been applied; they can be applied to keyframes prior to interpolation, or they can be applied to the entire morph weight set stream after interpolation, as explained below. A34, 5:16-20; A35, 7:19-32. An example of a post-processing rule might be to add a small amount of random noise to the morph weight set at each keyframe to introduce more variety into the character's expressions. A36, 9:41-45.

The patents' specification provides clear guidance on how families of correspondence and transition rules are created, as well as extensive examples for each category. *See* A35, 7:53-A36, 10:63. But the patents do not purport to list all such rules. That is because the morph weight set and transition times depend on the particular animation project; there is no single scientifically "correct" set of rules. For example, the contents of a morph weight set will vary depending on how many morph targets the artist has created for that model; it might contain 4 values or 30. Mouth shapes and facial expressions will also vary depending on the character. The same phoneme sequence may look very different when spoken by a robot, a cat, or a swamp monster. *See, e.g.*, A3573, :01-:26; A6121, Clip 3. Thus, as the patents explain, "[t]he rules of the present invention are extensible and

freeform in the sense that they may be created as desired and adapted to a wide variety of animation characters, situations, and products." A36, 9:23-26.

b. *Applying the Rules To Create Animation*

Once the rules have been established and the character model and visemes created, computer software does the rest. An essentially unlimited amount of video animation can be generated with no significant additional labor.

First, a specific sub-sequence of the time-aligned phonetic transcript of the dialogue is selected. A34, 5:49-50. This sub-sequence is like a moving "window" through the TAPT; there is a specific phoneme/time at its center, and several phonemes/times before and after it to provide context. The software evaluates the sub-sequence to see if any secondary correspondence rule applies. A34, 5:57-58. If one does, the designated morph weight set is applied. A34, 5:58-60. If no secondary rule applies, a morph weight set is applied according to the default rules. A34, 5:60-61. The software performs the same process for transition rules, looking for a secondary rule, and reverting to the default rule if none exists. A34, 5:61-65.

Applying the rules to each selected sub-sequence, in turn, generates a stream of morph weight sets and transition parameters over time; this stream is called the "intermediate file" or "intermediate stream." *See* A34, 5:65-67; A35, 7:10-11, 8:52-54; A37, 11:35. The morph weight sets in that stream act as keyframes, marking the important points of the speech. A34, 5:65-67; A35, 7:10-13; A36,

9:10-11. The software performs a pass to check if any pre-interpolation post-processing rules apply to the keyframes; if they do, they are applied at this time. A34, 5:67-6:2.

Once the keyframes are created, the animation software interpolates between the keyframe morph weight sets at the desired rate for the animation (*e.g.*, 30 images per second). A34, 6:3-4; A35, 7:15-18. Interpolation involves specifying morph weight sets on a frame-by-frame basis to provide a smooth transition between keyframes. A35, 7:13-16. The software then performs another pass to check if any post-interpolation post-processing rules apply; if so, those rules are applied to alter the values for both keyframes and interpolated frames. A34, 6:4-6; A35, 7:19-24. This yields the "final stream" of morph weight sets and transition parameters. A37, 11:40.

In the last step, the final stream of synthesized morph weight sets is applied to manipulate the neutral model (as described earlier) to automatically move the animated character's mouth and change its facial expression to match the recorded dialogue. A34, 6:20-24; A36, 9:18-21; A37, 11:44-47. That output is sent to a conventional computer animation display system for viewing, or to be integrated with other animation (such as the addition of background, other characters, etc.). A36, 9:18-20.

The end result is a video of an animated character speaking in sync with the pre-recorded dialogue, which can be used in movies, cartoons, video games, etc. A32, 2:49-50; A33, 3:22-23. The video tutorial's introduction contains several animation clips that were created using the patented invention. The contrast between the animation automatically produced by the patented method, and the "flappy," over-articulated animation produced by prior automated methods, must be seen to be appreciated. *Contrast* A3573, :01-:26; A6121, Clip 3 (animation produced by the invention), *with* A3573, 9:40-55, 17:51-18:12; A6121, Clip 2 (animation produced by the prior-art method).

III. The District Court's Decision

On September 22, 2014, the district court granted defendants judgment on the pleadings, holding that the patents cover abstract ideas that are not patent-eligible under 35 U.S.C. § 101.

The district court initially acknowledged that the § 101 analysis is governed by the two-step framework the Supreme Court established in *Mayo Collaborative Services v. Prometheus Labs, Inc.*, 132 S. Ct. 1289 (2012), and *Alice Corp. Pty. Ltd. v. CLS Bank International*, 134 S. Ct. 2347 (2014). A7. In the district court's view, however, "[d]escribing this as a two-step analysis may overstate the number of steps involved." *Id.* Whether a claim is abstract, the court stated, "may be more

like a one step test"—a sort of "'I know it when I see it'" decision. A8 (quoting *Jacobellis v. State of Ohio*, 378 U.S. 184, 197 (1964) (Stewart, J., concurring)).

After quoting Claim 1 of the '576 Patent and Claim 1 of the '278 Patent, the court stated that, "[f]acially, these claims do ***not*** seem directed to an abstract idea." A13 (emphasis added). To the contrary, "[t]hey are ***tangible***, each covering an approach to ***automated three-dimensional computer animation***, which is a ***specific technological process***." *Id.* (emphasis added). And the court rejected defendants' argument that the patents "claim a monopoly . . . on 'the idea that the human mouth looks a certain way while speaking particular sounds,' 'applied to the field of animation.'" *Id.* It noted that "the patents do ***not*** cover prior art methods of computer assisted, but non-automated, lip synchronization for three-dimensional computer animation." *Id.* (emphasis added). Indeed, defendants had urged that the patents "do not cover the automated methods of lip-synchronization for three-dimensional computer animation" that they use. A14. The court thus stated that, "[a]t first blush, it is . . . difficult to see how the claims might implicate the 'basic underlying concern that these patents tie up too much future use of' any abstract idea they apply." *Id.* (quoting *Mayo*, 132 S. Ct. at 1302).

Applying its own mode of analysis, however, the court stated that "it is not enough to view the claims in isolation"; rather, "when determining whether a patent contains an adequate inventive concept, the Court must factor out

conventional activity." A14. In the court's view, "conventional activity" encompasses ***anything*** found "***in the prior art***." A15 (emphasis added). The court thus created a claim chart, of the sort used in a § 103 obviousness analysis, listing each "Step" from the claims in one column, and supposedly "Admitted Prior Art" that corresponds to the step in the other column. *See* A17-18. The court found that, "while tangible, the steps of (1) using a timed phoneme transcript, (2) setting morph weight sets at keyframes, or (3) interpolating between keyframes" existed in the prior art, and thus could not be considered when deciding whether the claim was unpatentably abstract. A15.

Having read those "tangible" limitations out of the claims, the district court focused on what it considered the "point of novelty." A15; A17; *see also* A19. (The court did not address the Supreme Court's rejection of the point-of-novelty approach in *Diamond v. Diehr*, 450 U.S. 175, 188-90 (1981).) According to the court, "the point of novelty here is the idea of using rules, including timing rules, to automate the process of generating keyframes." A17. The court stated that the patents address the use of such rules "at the highest level of generality." A18. While the specification provides examples of rules, the claims are not limited to those examples. *Id.* The court therefore deemed the question before it as "whether the inclusion of that ***concept*** [of using rules] in the claims satisfies § 101 given (1) the prior art, and (2) the fact that the claims do not require any particular rules."

A17. The court held that it did not: "Because the claim purports to cover all such rules, in light of the prior art, the claim merely states 'an abstract idea while adding the words "apply it."'" A18 (quoting *Alice*, 134 S. Ct. at 2358). "The invention here may have been novel," the court stated, "but the claims are directed to an abstract idea." A19.

SUMMARY OF ARGUMENT

I. The claims here satisfy the two-step test for patent-eligibility under § 101 set forth by the Supreme Court in *Alice* and *Mayo*.

A. At step one, the claims are not "directed to" an abstract idea. The district court acknowledged as much, stating that "these claims do ***not*** seem directed to an abstract idea. They are ***tangible***, each covering an approach to ***automated three-dimensional computer animation***, which is a ***specific technological process***." A13 (emphasis added). That is correct. The claims cover a specific technological process for computer animation that yields a tangible result—a video of a 3-D character realistically speaking in sync with pre-recorded dialogue. The claims are not directed to a mere idea, having no particular concrete or tangible form.

These patents are not directed to a mathematical formula for calculating a number. Nor are they directed to a "business method" comprising ideas about

organizing human activity. They address an improvement to a specific technological process.

B. The § 101 analysis here thus should end at step one. But even if the Court were to assume the claims are directed to an abstract idea, they must be upheld if the claimed implementation "add[s] ***enough*** . . . to allow the processes they describe to qualify as patent-eligible processes that ***apply***" the putative abstract idea rather than seeking to monopolize the idea itself. *Mayo Collaborative Servs. v. Prometheus Labs, Inc.*, 132 S. Ct. 1289, 1297 (2012). A patent claim satisfies that test if it improves an existing technological process. And the claims here do just that. They recite a technological method that enables a computer to do something it could not do before—automatically produce realistic lip-synchronization animation.

The claims, moreover, extend only to a highly specific "application" of any underlying ideas. The patents do not simply say "use rules in animation." The claims cover ***only*** specific types of rules—those that "define output morph weight set stream ***as a function of phoneme sequence and time of said phoneme sequence***." A37, 11:30-32 (emphasis added). The patents thus are limited to a very specific computerized animation process—it must utilize "morph weight sets," and it must vary the 3-D character's expression based on the phoneme "sequence" (which sounds precede and follow the one being articulated), and

"time" (speed) of the speech. Those elements—which produced a "revolutionary" result in automated lip synchronization—render the invention a specific "application" of computer-implemented rules that is deserving of patent protection.

C. Because the claims recite only one specific means of computer animation among many, the claims do not implicate the fundamental pre-emption concern that undergirds the abstract-ideas exception. There are many "non-infringing ways" to use rules in lip-synchronization animation. A14. It is thus "difficult to see how the claims might implicate the 'basic underlying concern that these patents tie up too much future use of' any abstract idea they apply." *Id.* (quoting *Mayo*, 132 S. Ct. at 1302).

D. The claims' patent-eligibility is further confirmed under the machine-or-transformation test. The claims transform a phonetic transcript of dialogue and static character models into a realistic video of the character speaking, synchronized to pre-recorded dialogue, that can be watched on a screen. They transform a general-purpose computer into a specific-purpose machine that does something computers could not do before—automatically create realistic lip-synchronization animation.

II. The district court found the patents abstract only after departing from the *Alice/Mayo* test and applying a novel § 101 test of its own devising. But the court's analysis has already been roundly criticized, *see Cal. Inst. of Tech. v.*

Hughes Commc'ns Inc., — F. Supp. 3d —, 2014 WL 5661290, at *11 (C.D. Cal. Nov. 3, 2014), and with reason.

A. First, at *Alice/Mayo* step one, the court excluded from its analysis any limitation with a basis in the prior art, declaring the patents' supposed "point of novelty" to be abstract. But the Supreme Court expressly rejected that approach in *Diamond v. Diehr*, 450 U.S. 175 (1981), holding that "[t]he 'novelty' of any element or steps in a process . . . is of no relevance" in determining whether a claim is directed to an abstract idea. *Id.* at 188-89. To the extent purely "conventional activity" may sometimes be discounted, that factor is considered only in the context of the ***second step*** of the *Alice/Mayo* test. And the district court interpreted "conventional activity" to mean that ***any step with a basis in the prior art*** must be disregarded in the § 101 analysis. *See* A14-19. But neither *Mayo* nor any other precedent defines "conventional activity" to include ***everything*** in the prior art.

B. The district court made erroneous findings even within its own faulty "point of novelty" framework, misconstruing the scope of the prior art. And the court ultimately held that the patents' use of rules at the supposed "point of novelty" is an "abstract idea" because the "concepts are specified at the highest level of generality." A18. But the rules are not claimed at the highest level of generality. The specific types of rules are identified: those based on "phoneme sequence" and "timing of phoneme sequence." Claiming those categories of rules, rather than

reciting every example, is accepted patent practice. It does not render the claims abstract.

C. The district court's analysis would endanger not just patents relating to computer animation, but all software patents. As one judge explained in criticizing the decision below, "most inventions today build on what is known in the art, and an improvement to software will almost inevitably be an algorithm or concept which, when viewed in isolation, will seem abstract." *Caltech*, 2014 WL 5661290, at *11. The district court's point-of-novelty analysis "would likely render all software patents ineligible." *Id.*

STANDARD OF REVIEW

This Court reviews *de novo* a district court's determination of patent-eligibility under 35 U.S.C. § 101. *DDR Holdings, LLC v. Hotels.com, L.P.*, 773 F.3d 1245, 1255 (Fed. Cir. 2014).

ARGUMENT

The claims at issue are directed to a "method for automatically animating lip synchronization and facial expression of three-dimensional characters" in computer animation. A37, 11:27-28. In the prior art, high-quality animation could be produced manually, by a human artist sitting down at a computer and creating the images at each keyframe. That method, however, was time-consuming and expensive. More automated methods also existed, but the quality was poor and un-

convincing. The claimed method here overcomes those limitations and provides the best of both worlds—it automatically moves the face of a computer-generated model so that it produces realistic video of the character speaking pre-recorded dialogue. *See* pp. 18-25, *supra*. Defendants themselves hailed the invention as "revolutionary." A4995.

The district court's holding that the claims do not cover patent-eligible subject matter under § 101 cannot stand. The district court admitted that, "[f]acially, ***these claims do not seem directed to an abstract idea***. ***They are tangible***, each covering an approach to automated three-dimensional computer animation, which is ***a specific technological process***." A13 (emphasis added). For that reason, the claims independently satisfy each prong of the two-part test for patent-eligibility the Supreme Court adopted in *Alice* and *Mayo*. They likewise satisfy this Court's machine-or-transformation test, which remains "a useful and important clue" that the claims are patent-eligible. *See Bilski v. Kappos*, 561 U.S. 593, 604 (2010). And it is clear that the patents do not seek to monopolize anything remotely resembling the "building blocks of human ingenuity" or "the basic tools of scientific and technological work." *Alice*, 134 S. Ct. at 2354. There are numerous ways to perform lip-synchronization animation, automated lip synchronization, and even automated, rules-based lip synchronization, that do not infringe the patents.

The district court found the patents abstract only after applying a novel framework that resembled a § 103 obviousness analysis. Reading out of the claims any limitation with a basis in the prior art, the court attempted to locate the supposed "point of novelty." A17. But the Supreme Court has expressly rejected that approach. And the district court made erroneous findings regarding the patented invention and the prior art, even within its own framework.

As another judge from the Central District of California recently explained in directly criticizing the decision below,

> ***[I]t is difficult to imagine any software patent that survives under [the decision below's] approach***—most inventions today build on what is known in the art, and an improvement to software will almost inevitably be an algorithm or concept which, when viewed in isolation, will seem abstract. ***This analysis would likely render all software patents ineligible*** . . . .

Cal. Inst. of Tech. v. Hughes Commc'ns Inc., — F. Supp. 3d —, 2014 WL 5661290, at *11 (C.D. Cal. Nov. 3, 2014) (emphasis added). The decision below should be reversed.

I. The Claims Are Not Unpatentably Abstract Under § 101

Section 101 defines patentable subject matter as "any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof." 35 U.S.C. § 101. But "laws of nature, natural phenomena, and abstract ideas" are not eligible. *Diamond v. Diehr*, 450 U.S. 175, 185 (1981).

To determine whether a patent covers an abstract idea outside § 101, the court first considers whether the claims are "directed to" an abstract idea. *Mayo*, 132 S. Ct. at 1296-97. If they are, the court considers whether the claims' elements "add ***enough*** to their statements of the [abstract idea] to allow the processes they describe to qualify as patent-eligible processes that ***apply*** [the abstract idea]." *Id.* at 1297. The Supreme Court has "described step two of this analysis as a search for an 'inventive concept'—*i.e.*, an element or combination of elements that is 'sufficient to ensure that the patent in practice amounts to significantly more than a patent upon the [ineligible concept] itself.'" *Alice*, 134 S. Ct. at 2355 (quoting *Mayo*, 132 S. Ct. at 1294) (alteration in original).

Alice clarified that the abstract-ideas exception does not apply if the invention "solve[s] a technological problem in 'conventional industry practice,'" "improve[s] an existing technological process," or otherwise "effect[s] an improvement in any other technology or technical field." 134 S. Ct. at 2358, 2359. While the Court did not hold that an invention ***must*** represent a technological advance to be patent-eligible under § 101, *Alice* indicates that a claim that ***does*** represent such an advance is patent-eligible.

The claims here independently satisfy ***each*** step of the *Alice/Mayo* test.

A. Step One: The Claims Are Not "Directed To" an Abstract Idea

1. *The Claims Are "Directed To" a Technological Process That Produces Tangible Results*

At the first *Alice/Mayo* step, the court must make a threshold determination "whether the claims at issue are directed to a patent-ineligible" abstract idea. *Alice*, 134 S. Ct. at 2355. Here, they are not.

Representative Claim 1 of the '576 Patent recites:

> A method for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:
>
> obtaining a first set of rules that define output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence;
>
> obtaining a timed data file of phonemes having a plurality of subsequences;
>
> generating an intermediate stream of output morph weight sets and a plurality of transition parameters between two adjacent morph weight sets by evaluating said plurality of sub-sequences against said first set of rules;
>
> generating a final stream of output morph weight sets at a desired frame rate from said intermediate stream of output morph weight sets and said plurality of transition parameters; and
>
> applying said final stream of output morph weight sets to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters.

A37, 11:27-47.

The claims expressly state their purpose: "automatically animating lip synchronization and facial expression of three-dimensional characters." A37, 11:27-

28. And the claimed method generates a tangible product—"lip synchronization and facial expression control of . . . animated characters." A37, 11:45-47.

Every claim element is in service of, and necessary to, the recited method of automating lip synchronization and facial expression in 3-D computer animation. They do not merely recite "use rules in animation," or "use rules on a computer to produce animation." They require particular types of rules—those that analyze a phonetic transcript as a "function of ***phoneme sequence*** and ***time of said phoneme sequence***," A37, 11:31-32 (emphasis added)—as part of a specific, "integrated method," A32, 2:40, for using "morph weight sets" to generate video of a character speaking, A37, 11:44-47. No limitation is "plainly . . . divisible" from the other elements as a stand-alone abstract concept. *DDR Holdings*, 773 F.3d at 1256. The district court thus was correct when it acknowledged that, "[f]acially, ***these claims do not seem directed to an abstract idea***," and that "***[t]hey are tangible***, each covering an approach to automated three-dimensional computer animation, which is ***a specific technological process***." A13 (emphasis added).

Indeed, the entire field of 3-D computer-generated animation is ***inherently*** technological and tangible. Even using prior-art methods, the critical steps—creating the neutral character model and morph targets; blending models and morph weight sets in determining the keyframes; and interpolating between keyframes—are performed using special animation software on computers. *See*

pp. 13-15, *supra*. The method for performing that process "automatically" here, A37, 11:27, "embodied in computer software for use with a computer," A32, 2:40-41, is likewise inherently technological. It is no mere "idea, having no particular concrete or tangible form." *Ultramercial, Inc. v. Hulu, LLC*, 772 F.3d 709, 715 (Fed. Cir. 2014).

The purpose of the claims, moreover, is to ***make something tangible***. The method produces a video of a 3-D character speaking the recorded audio—video you can see when watching a cartoon or playing a video game. That tangible output is an element of the claim. *See* A37, 11:44-47 ("applying said final stream of output morph weight sets to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters"). It is hard to see how anyone could watch video clips of animation produced by the patented method, *see* A3573, :01-26; A6121, Clip 3, and conclude that the claims are directed to a mere "abstraction," *Ultramercial*, 772 F.3d at 715.

2. *Supreme Court Precedent Confirms That the Claims Are Not "Directed To" an Abstract Idea*

The Supreme Court has recognized two categories of claims that implicate the abstract-ideas exception. The first concerns claims covering algorithms, in the form of mathematical formulas, that are used for calculating numbers. In *Parker v. Flook*, 437 U.S. 584 (1978), for example, the Court held that a claim covering a formula for calculating "alarm limits"—which were simply "a number"—was an

unpatentable abstract idea. *Id.* at 585. Similarly, in *Gottschalk v. Benson*, 409 U.S. 63 (1972), the Court held that a claim to a mathematical formula for converting binary-coded decimals into pure binary numerals was unpatentably abstract. *Id.* at 64. Second, the Court has found so-called "business methods"—essentially ideas about "fundamental economic practice[s]" and "organizing human activity"—to be abstract. *Alice*, 134 S. Ct. at 2356-57. In *Alice*, the Court invalidated claims directed to the business method of "intermediated settlement." *Id.* And in *Bilski*, the Court held that claims directed to "hedging risk" were abstract ideas. 561 U.S. at 609. The claims here do not remotely fit within those categories.

Instead, they are like the claim in *Diehr*, 450 U.S. at 184, which the Supreme Court held was ***not*** directed to an abstract idea. The claim there was for a "method of operating a rubber-molding press for precision molded compounds with the aid of a digital computer." *Id.* at 179 n.5. It recited the use of a mathematical formula, the "Arrhenius equation," as part of a "step-by-step method" for curing rubber. *Id.* at 179 & n.5, 184. The Court explained that "Arrhenius' equation is not patentable in isolation." *Id.* at 188. But the claim was not directed to "patent[ing] [that] mathematical formula." *Id.* at 187. Instead, it sought "patent protection for a process of curing synthetic rubber." *Id.* The Court stated that "[i]ndustrial processes

such as this are the types which have historically been eligible to receive the protection of our patent laws." *Id.* at 184.

As in *Diehr*, the claims here do not seek to patent a "mathematical formula" or any other abstract concept. Instead, they cover a specific, step-by-step process—implemented through software—for automatically producing a video of a 3-D computer-generated character that speaks in sync with pre-recorded audio. No less than the rubber-curing method in *Diehr*, that is a specific technological process that produces a tangible result. It, too, should be "eligible to receive the protection of our patent laws." *Diehr*, 450 U.S. at 184.

3. *The Claims Are Not "Directed To" an Abstract Idea Under This Court's Precedent*

This Court's post-*Alice* cases are to the same effect. In *Digitech Image Technologies, LLC v. Electronics for Imaging, Inc.*, 758 F.3d 1344 (Fed. Cir. 2014), the Court found claims directed to a "device profile," which was a formula for combining two data sets into one, to be ineligible under § 101. *Id.* at 1351. Like the claims in *Flook*, it was simply a means of calculating numbers. *Id.* And in *Ultramercial*, 772 F.3d at 715, and *buySAFE, Inc. v. Google, Inc.*, 765 F.3d 1350, 1355 (Fed. Cir. 2014), the Court invalidated patents directed to business methods—"using advertising as an exchange or currency," and creating a "transaction performance guaranty," respectively—that were not distinguishable from the claims the Supreme Court invalidated in *Alice* and *Bilski*. *See also Planet*

Bingo, LLC v. VKGS LLC, 576 F. App'x 1005, 1008 (Fed. Cir. 2014) (invalidating claims for "managing a game of bingo" as "similar to the kind of 'organizing human activity' at issue in *Alice*").

The claims here are quite different. They more closely resemble the patent in *DDR Holdings*. The patent there addressed a problem "particular to the Internet"—how a host website can retain visitors when the visitor clicks on a link to a third-party merchant's advertisement. 773 F.3d at 1257. It claimed a system that generates a "hybrid" website that retains the "look and feel" of the host's website, while allowing the visitor to buy products from the third-party merchant without actually entering the merchant's website. *Id.* at 1257-58.

Rejecting the contention that the claims sought to patent abstract ideas, the Court observed that the claims did not fall within the categories previously found to implicate the abstract-ideas exception: "[The] claims do not recite a mathematical algorithm. Nor do they recite a fundamental economic or longstanding commercial practice." 773 F.3d at 1257. While the claims implicated commerce, the Court found, "the claimed solution is necessarily rooted in computer technology in order to overcome a problem specifically arising in the realm of computer networks." *Id.* And while the claims at issue were not "technologically complex," they were nevertheless technological: They "specify how interactions with the Internet are manipulated to yield a desired result" when clicking a hyper-

link. *Id.* at 1258-59. The claims were "different enough in substance" from claims in prior cases that "broadly and generically claim[ed] 'use of the Internet' to perform an abstract business practice" to be patent-eligible. *Id.* at 1258.[2]

As in *DDR Holdings*, the claims here are "necessarily rooted in computer technology in order to overcome a problem specifically arising in the realm of computer[s]." 773 F.3d at 1257. They provide a method for getting a computer to automatically generate video of a 3-D animated character speaking in sync with pre-recorded dialogue—without requiring an artist's constant intermediation, or yielding the unrealistic results of prior automated methods. *See* pp. 18-25, *supra*. Like the claims in *DDR Holdings*, they are patent-eligible because they constitute a technological advance that is sufficiently "unlike the claims in *Alice*" and other cases "that were found to be 'directed to' little more than an abstract concept." 773 F.3d at 1259. Indeed, the claims here are more clearly patent-eligible than those in *DDR Holdings*. Unlike *DDR Holdings*, there is no conceivable argument that the claims are merely "entrepreneurial" rather than "technological."

[2] Judge Mayer dissented. In his view, the patents were directed to the idea "that an online merchant's sales can be increased if two web pages have the same 'look and feel.'" 773 F.3d at 1263. Because the "solution they offer is an entrepreneurial, rather than a technological, one," Judge Mayer would have found the claims unpatentably abstract. *Id.* at 1265.

B. Step Two: The Claims Recite a Patent-Eligible *Application* of Any Idea

Step two of the *Alice/Mayo* framework assumes the court has found that the patent claims are directed to an abstract idea at step one. Because the claims here are not directed to an abstract idea at all, the Court need go no further. But even if the Court were to assume the patent ***is*** directed to an abstract idea, the implementation here "add[s] ***enough*** . . . to allow the [claimed] processes . . . to qualify as patent-eligible processes that ***apply***" any putative abstract idea. *Mayo*, 132 S. Ct. at 1297.

1. *The Claims Are a Patent-Eligible Improvement to a Technological Process*

Defendants urged that the claims here are directed to "the abstract idea of rules-based synchronization of animated mouth movement." D.Ct.Dkt.#338, at 12. In *Alice*, however, the Supreme Court indicated that a claim represents a patent-eligible ***application*** of an idea if it "effect[s] an improvement in any other technology or technical field." 134 S. Ct. at 2359. The invention here provides just such an improvement in the technological field of 3-D computer-generated lip-synchronization animation.

As explained above, the field of the invention—3-D animation—is inherently technological. All of the character models are created and manipulated using special software on computers. *See* pp. 4-15, *supra*. The problem the patents

solve is also a technological one: How can one improve animation software so that it can analyze a phonetic transcript of spoken dialogue and automatically manipulate a computer-generated facial model to create a video of the character that realistically looks like it is speaking the audio dialogue?

In the prior art, realistic animation was achievable only if a human sat at a computer and painstakingly set the morph weight set at each keyframe, manually manipulating sliders in animation software. *See* pp. 13-14, *supra*. That process was "laborious," "lengthy," and expensive. A32, 1:21. Other prior-art methods utilized software to automatically produce lip-synchronization animation, but the results were unrealistic. *See* pp. 14-15, *supra*.

Here, the patents teach Rosenfeld's critical insight about how a computer can automatically produce realistic lip-synchronization from a phonetic transcript if it is programmed to take into account not just phonemes, but also ***context*** and ***pace*** of speech. *See* A4950 ¶12; A4953 ¶20. The claims' implementation of that insight—programming animation software with "rules that define output morph weight set stream as a ***function of phoneme sequence and time of said phoneme sequence***," A37, 11:30-32 (emphasis added)—is technological. It is also clearly a "functional and palpable" improvement, *Research Corp. Techs., Inc. v. Microsoft Corp.*, 627 F.3d 859, 868 (Fed. Cir. 2010), because it enables a computer to do

something it could not do before: automatically produce high-quality lip synchronization animation from a phonetic transcript.

Comparing a sample of animation produced using prior automated methods, *see* A3573, 9:40-55, 17:51-18:12; A6121, Clip 2, with clips of animation produced using the patented method, *see* A3573, :01-:26; A6121, Clip 3, illustrates precisely how the method "improve[s] an existing technological process," *Alice*, 134 S. Ct. at 2358. It is the type of "new and useful advance[] in technology" the patent system was designed to encourage and protect. *Pfaff v. Wells Elecs., Inc.*, 525 U.S. 55, 63 (1998).

2. *The Inventive Concepts Recited in the Claims Represent a Specific Application*

Even apart from meeting *Alice*'s "technological improvement" standard, the claims satisfy step two of the *Alice/Mayo* analysis because they reflect an "'inventive concept'—*i.e.*, an element or combination of elements that is 'sufficient to ensure that the patent in practice amounts to significantly more than a patent upon'" any purported abstract idea "'itself.'" *Alice*, 134 S. Ct. at 2355 (quoting *Mayo*, 132 S. Ct. at 1294). Indeed, the claims contain several such inventive concepts. They recite a method that employs specific ***types*** of rules to produce high-quality, computer-generated facial animation. And the claims provide a specific technological ***way*** of using those rules to generate the animation.

a. The claims cover ***only*** the use of "rules that define output morph weight set stream ***as a function of phoneme sequence and time of said phoneme sequence***." A37, 11:30-32 (emphasis added). That reflects a central "inventive concept"—Rosenfeld's insight that, to create high-quality animation automatically, one needs ***only*** to set rules enabling the animation software to account for ***those specific parameters***. *See* A4950 ¶12; A4953 ¶20. The rules adjust for the fact that a phoneme may look different when spoken depending on the phonemes preceding and/or following it, and for the fact that a viseme's expression is affected by how quickly the character speaks. That drastically limits the claims' scope. It renders the invention a specific "***application***" of the use of computer-implemented rules that is "deserving of patent protection." *Diehr*, 450 U.S. at 187 (emphasis added).

The claims do not recite individual rules regarding "the phoneme sequence" and "time of said phoneme sequence." But that is immaterial. The rules will vary depending on the taste of the animator and the particular character she has created—a swamp monster will use different rules than a tight-lipped cat.[3] It is

[3] The "extensible and freeform" nature of the claimed rules is thus both a virtue and a necessity, because they must be "adapted to a wide variety of animation characters, situations, and products." A36, 9:23-26. The patent does, however, explain in detail how to create such rules in the context of the claimed method, describing: (1) how they break down into correspondence rules and transition rules; (2) how the rules can further be grouped into default, secondary, and post-

settled law that the patentee may write claims at a level that encompasses the genus of the invention, rather than reciting every species. *See AbbVie Deutschland GmbH & Co., KG v. Janssen Biotech, Inc.*, 759 F.3d 1285, 1299 (Fed. Cir. 2014).

Here, the claims recite "rules that define output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence." A37, 11:30-32. If that sweeps more broadly than the patents' teachings can justify (and it does not), that would be an issue under § 112's "written description" requirement, not § 101. *See Research Corp.*, 627 F.3d at 869; *AbbVie*, 759 F.3d at 1299. Or, if the claims are overly broad, they risk covering prior art under § 102 or being obvious under § 103. But claiming the category of rules based on "phoneme sequence" and "timing of phoneme sequence," as opposed to reciting every example of those rules, does not render the claims "so manifestly abstract as to override the statutory language of section 101." *Research Corp.*, 627 F.3d at 869. Neither defendants nor the district court explained how software ***could*** be patentable if specific rules had to be claimed.

b. There are, moreover, "additional elements" of the claim that, taken "as an ordered combination," further "transform" the invention "into a patent-eligible application." *Alice*, 134 S. Ct. at 2355, 2357. The claims require that the method analyze a time-aligned phonetic transcript of the dialogue and "automati-

---

processing rules to achieve different functions; and (3) specific examples of each type of rule. *See* pp. 20-23, *supra*.

cally" generate a "stream" of "morph weight sets" and "transition parameters" based on that transcript in light of the established rules regarding the sequence and timing of phonemes. A37, 11:27-39. The "stream of output morph weight sets" is manipulated and ultimately applied "to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters." A37, 11:40-47. The claims' recitation of "morph weight sets" thus limits the claims to one specific, technological process for producing lip-synchronization animation. *See* pp. 23-25, *supra*.

The record shows that the claims' recitation of "morph weight sets" cannot be written off as "purely conventional" activity. *Alice*, 134 S. Ct. at 2358. For example, defendant Naughty Dog petitioned the PTO's Patent Trial and Appeal Board to institute an *inter partes* review of the claims at issue, claiming they were anticipated or obvious. *See* A4363-64. In denying review, the PTAB found that Naughty Dog had failed to show that at least three of the cited references were anticipatory because they did ***not*** utilize morph weight sets. A4377; A4380; A4382. And in the district court below, defendants argued that ***their*** methods of automated animation do not infringe because they do not utilize morph weight sets. *See* D.Ct.Dkt.#329, at 11-12; A14. If the presence or absence of the "morph weight set" limitation is potentially dispositive of both validity and infringement, by definition it "add[s]" something "of practical significance to the underlying ab-

stract idea." *Ultramercial*, 772 F.3d at 715, 716. It necessarily limits the scope of the claims in a meaningful way, "transform[ing]" the idea of using rules in lip-synchronization animation into a specific, technological "application." *Mayo*, 132 S. Ct. at 1294.

C. The Claims Do Not Preempt the Idea of Using Rules in Lip-Synchronization Animation

For the reasons above, the claims do not implicate the fundamental "pre-emption concern that undergirds" the abstract-ideas exception. *Alice*, 134 S. Ct. at 2358. The claims do not, as defendants assert, seek to "patent[ ] the abstract idea of using rules itself," and thus "wholly pre-empt the use of ***any*** rules" relating to lip-synchronization animation. D.Ct.Dkt.#338, at 13 (emphasis in original). To the contrary, an animator may utilize any rules he wishes. He infringes only if he uses software that "define[s] output morph weight set stream as ***a function of phoneme sequence*** and ***time of said phoneme sequence***." A37, 11:30-32 (emphasis added).

For example, one prior-art automated method used "rules" of a sort to generate facial animation: Whenever a phoneme appeared in the TAPT, the corresponding viseme was used (*i.e.*, 100% of that viseme) to the exclusion of all others. *See* pp. 14-15, *supra*. That rule did not infringe because it does not take into account ***context*** and ***timing*** of a phoneme sequence. And widely used facial-capture technologies employ rules in the form of a "solver" to automatically

generate morph weight sets for animation. *See* Barbara Robertson, *Big Moves*, Computer Graphics World (Nov. 2006), *available at* http://www.cgw.com/Publications/CGW/2006/Volume-29-Issue-11-Nov-2006-/Big-Moves.aspx. Those rules, however, are based not on the timing and sequence of phonemes, but on sampling differences in the relative positions of an actor's face. *See id.* That is another example of the many different types of rules that can be devised for use in animation that are not covered by the claims.

The additional limitations requiring that the method "automatically" produce "streams" of "morph weight sets" from a TAPT further limit the claims' pre-emptive scope. They exclude a vast swath of animation techniques, ***even if*** they were to employ rules based on "function of phoneme sequence and time of said phoneme sequence." A37, 11:26-47. Rules could guide an artist when producing hand-drawn, two-dimensional animation for movies like *The Lion King* and cartoons like *Dora the Explorer*. But that would not infringe because the process does not utilize morph weight sets. Likewise, it is possible to define rules that an artist would use in prior-art techniques for manually producing 3-D computer-generated animation features, such as Pixar's *Toy Story* movies. But that process does not automatically produce morph weight sets from the TAPT. *See* pp. 13-14, *supra*. Several other prior-art references Naughty Dog cited in its request to initiate an IPR did not utilize morph weight sets at all. *See* p. 48, *supra*. And

defendants have argued that the methods they use to automate animation in their video games do not use morph weight sets. *See id.* As the district court recognized, given the many "noninfringing ways" to use rules in lip-synchronization animation, it is "difficult to see how the claims might implicate the 'basic underlying concern that these patents tie up too much future use of' any abstract idea they apply." A14 (quoting *Mayo*, 132 S. Ct. at 1302).

The patents thus in no way "impede innovation." *Alice*, 134 S. Ct. at 2354. They spur innovation. Rosenfeld came up with a new way of using morph weight sets and specific rule types to generate improved automated lip-sync animation. If an animator uses software that employs other rule types yielding better animation than Rosenfeld's invention, the claims do not cover it. Or an animator can use the same rules in an improvement that does not involve morph weight sets; the claims would not cover that, either. The "building blocks of human ingenuity" and "basic tools of scientific and technological work" remain free to all. *Id.* The patents' scope is no larger than "the underlying discovery could reasonably justify." *Mayo*, 132 S. Ct. at 1301.

D. The Claims "Transform" Static Inputs Into Video Using a Special-Purpose "Machine"

Before the Supreme Court established the *Alice/Mayo* framework, this Court used "the machine-or-transformation test" when "determining patent eligibility of a process under § 101." *In re Bilski*, 545 F.3d 943, 956 (Fed. Cir. 2008) (en banc),

aff'd sub nom. Bilski v. Kappos, 561 U.S. 593 (2010). Under that test, a process is patent-eligible if: "(1) it is tied to a particular machine or apparatus, or (2) it transforms a particular article into a different state or thing." *Id.* at 954. Although it "is not the sole test" for patent-eligibility, the machine-or-transformation test remains "a useful and important clue." *Bilski*, 561 U.S. at 604. The claims here satisfy both prongs of that test.

First, the claims are transformative. They change static, computer-generated models, based on a transcript, to create realistic, talking, 3-D animated characters. In particular, the method starts with a time-aligned phonetic transcript of prerecorded dialogue, A37, 11:34-35, which is "tangible," A15, and the neutral model and morph targets, which are also tangible. It then applies the specified steps to transform those into a "stream of output morph weight sets and a plurality of transition parameters." A37, 11:36-39. It further manipulates the output morph weight set stream, and eventually "appl[ies] said final stream of output morph weight sets to a sequence of animated characters ***to produce lip synchronization and facial expression control of said animated characters***." A37, 11:26-47 (emphasis added). Or, as the district court put it during claim construction, a "morph weight set" is a "set of values . . . that, when applied, ***transform*** the neutral model to some desired state." D.Ct.Dkt.#298-1, at 9 (emphasis added). In other words, the concrete inputs are transformed into a video of the character speaking in

sync with the dialogue that can be watched on a screen. *See* A36, 9:18-21. Because the invention "transform[s] . . . raw data into a particular visual depiction of a physical object on a display," that is "sufficient to render" the claimed process "patent-eligible." *Bilski*, 545 F.3d at 963; *see In re Abele*, 684 F.2d 902 (C.C.P.A. 1982) (upholding invention that takes X-ray attenuation data and displays it in visual form on a screen).

The claims, which cover an "integrated method embodied in computer software for use with a computer," A32, 2:40, are also tied to a special-purpose machine. They transform a general-purpose computer into a special-purpose machine that does something computers could not do before—automatically create realistic lip-synchronization animation. The method's use of a new, "specific machine to produce a useful, concrete, and tangible result" further supports patent-eligibility. *In re Alappat*, 33 F.3d 1526, 1544-45 (Fed. Cir. 1994) (en banc).

II. The District Court's §101 Analysis Violates Supreme Court Precedent and Threatens All Software Patents

The district court agreed that, at step one of the *Alice/Mayo* analysis, "[f]acially, these claims do not seem directed to an abstract idea. They are tangible, each covering an approach to automated three-dimensional computer animation, which is a specific technological process." A13. That should have ended the inquiry. Instead, the court improvised a new analytical framework, stating that, "the claims must be evaluated in the context of the prior art." A14. Using a claim

chart, the district court disregarded every element with a basis in the prior art. A17-18. It therefore read admittedly "tangible" steps out of the claims, A15, and focused on what it considered the "point of novelty"—"the idea of using rules, including timing rules, to automate the process of generating keyframes." A17.

Several months after the decision issued, Judge Pfaelzer—who sits in the same courthouse as Judge Wu—issued *Caltech*, a § 101 decision that cogently explains the flaws in the decision below. As Judge Pfaelzer recognized: (1) the analysis below improperly conflates steps 1 and 2 of the *Alice/Mayo* analysis; (2) it invokes a "point of novelty" approach the Supreme Court has rejected; and (3) it would threaten all software patents. 2014 WL 5661290, at *10-11. The district court's analysis, moreover, departs from the text of the claims here and misconstrues the prior art.

A. The District Court's Analysis Violates Supreme Court Precedent

1. *The Supreme Court Has Rejected the District Court's "Point of Novelty" Approach*

In *Diehr*, the Supreme Court expressly rejected the "point of novelty" approach the district court applied here. The Court stated that, in "determining the eligibility of [the patentees'] claimed process for patent protection under § 101, their claims must be considered as a whole." 450 U.S. at 188. "It is inappropriate to dissect the claims into old and new elements and then to ignore the presence of the old elements in the analysis." *Id.* The Court explained that "[t]he ***'novelty' of***

any element or steps in a process . . . ***is of no relevance*** in determining whether the subject matter of a claim falls within the § 101 categories of possibly patentable subject matter." *Id.* at 188-89 (emphasis added). The district court here did precisely what *Diehr* prohibits—it dissected the claims "into old and new elements" and focused ***solely*** on the "point of novelty" in determining whether the claims' subject matter is patent-eligible in the first instance.

The district court purported to base its analysis on *Mayo*, which stated that "well-understood, routine, conventional activity . . . is normally not sufficient to transform an unpatentable law of nature into a patent-eligible application of such a law." 132 S. Ct. at 1298; A14. But as Judge Pfaelzer explained, that "conflates step one and step two of *Mayo*." *Caltech*, 2014 WL 5661290, at *11. The district court purported to do that at step one because, in its view, "the two-step test may be more like a one step test." A8.

That was error. As this Court has explained, "any novelty in implementation of the idea is a factor to be considered ***only in the second step*** of the *Alice* analysis." *Ultramercial*, 772 F.3d at 715 (emphasis added). If a court finds that claims are directed to an abstract idea in the first instance, the presence of merely routine or conventional activity may not "add ***enough*** to their statements of the [abstract idea] to allow the processes they describe to qualify as patent-eligible processes that ***apply*** [the abstract idea]." *Mayo*, 132 S. Ct. at 1297. But if a claim

is, as the district court found here, directed to a "tangible," "specific technological process" in the first instance, A13, neither *Mayo* nor any other case suggests that a court should then strip the claims down to the "point of novelty" to invalidate it under § 101. Indeed, *Diehr* expressly prohibits that.

2. *The District Court Misconstrued the Supreme Court's Exclusion of "Purely Conventional Activity"*

The district court also fatally misconstrued the ***scope*** of the "conventional activity" exclusion. *Mayo* explains that "well-understood, routine, conventional activity" may not suffice to "transform an unpatentable law of nature into a patent-eligible application of such a law." 132 S. Ct. at 1298. The district court, however, interpreted that to mean that ***any step with a basis in the prior art*** must be disregarded. *See* A14-19. As Judge Pfaelzer explained in criticizing the decision below, "neither *Mayo* nor any other precedent defines conventional elements to include ***everything*** found in prior art." *Caltech*, 2014 WL 5661290, at *11.

The point of the "conventional activity" rule is that, where a practice is ubiquitous among those "who work in the field," reciting that activity will not meaningfully narrow the scope of the claim. *Mayo*, 132 S. Ct. at 1298. Put differently, where a court determines that a claim recites an abstract idea accompanied solely by steps so "conventional" that any practitioner would assume them necessary, that amounts to nothing more than "simply stat[ing] the [abstract idea] while adding the words 'apply it.'" *Id.* at 1294.

Rather than evaluate whether the claims merely recite an idea along with purely "conventional" activity, the district court created a claim chart of the sort that would be used in a § 103 obviousness analysis and disregarded every step of the claims that, in its view, was "admitted" to be in the prior art. *See* A17-18. It never asked the fundamental question whether the steps were so "conventional" that their recitation failed to meaningfully limit the claims' scope to a particular "application." A17.[4]

Stripping out every claim element with a supposed basis in prior art—as opposed to disregarding truly "conventional" activity—distorted the outcome. For example, the district court considered "setting morph weight sets at keyframes" to be a "tangible" step in the claims. A15. It did not consider that step in the § 101 analysis, however, because it had a basis in the prior art. A15, A17-18. But even if that step was in the prior art, it was not so "conventional" that it failed to limit the scope of the claims. As explained above, when defendant Naughty Dog sought *inter partes* review of the patents, the PTAB found that a number of prior-art methods were not anticipatory because they did not use morph weight sets. *See* p. 48, *supra*. And defendants have urged that they do not infringe because they do

[4] Nor did the court suggest that the steps were mere "generic computer" functions of the sort courts have disregarded in the § 101 analysis. *See Alice*, 134 S. Ct. at 2358-60 ("wholly generic computer implementation" using only "basic calculation, storage, and transmission functions"); *buySAFE*, 765 F.3d at 1355 (generic computer function of "receiving and sending information over networks").

not utilize morph weight sets. *See id.* Thus, use of morph weight sets is not so "conventional" that it should be summarily disregarded under *Mayo*. Especially when considered as part of the "ordered combination" that makes up the claims here—a combination that requires morph weight sets and specified rule parameters—it helps "transform" any "idea" of using rules into a "patent-eligible application." *Alice*, 134 S. Ct. at 2355, 2357.

B. The District Court's Conclusions Are Erroneous

The district court's analysis also rests on faulty conclusions about what was purely conventional or in the prior art. And the putative "point of novelty" it identified was anything but abstract.

1. *The District Court's Assertion That Rules Relating to Phoneme Sequence Were in the Prior Art Is Unsupported*

The district court misinterpreted the prior art. Most critically, the patents claim "rules that define output morph weight set stream ***as a function of phoneme sequence and time of said phoneme sequence***." A37, 11:30-32 (emphasis added). The district court found that "[r]ules for defining morph weight sets ***as a function of timing***" were not disclosed in the prior art. A17 (emphasis added). But it stated that the patents "admitted" that "[r]ules for defining morph weight sets ***as a function of phoneme sequence*** are disclosed as within the prior art." A17 (emphasis added).

That was clear error. The district court gave no explanation; it just cited the ’576 Patent at 1:44-2:28 (A32). That portion of the specification describes the prior-art practice of an artist manually applying morph weights. A32, 1:44-2:15. It notes that morph weights may be applied to visemes, and that visemes can be blended. A32, 2:16-2:28. But it does not discuss rules. It makes no mention of phoneme sequence. And it certainly does not “admit” that the prior art utilized “rules that define output morph weight set stream as a function of phoneme sequence.” A37, 11:30-32. The district court erred in excluding that critical component of the invention from the § 101 analysis. *See ActiveVideo Networks, Inc. v. Verizon Commc’ns, Inc.*, 694 F.3d 1312, 1340 (Fed. Cir. 2012). It is not prior art, much less a purely conventional step.

2. *The Claimed Sequence- and Timing-Based Rules Are Not Unpatentably Abstract Even When Isolated from Other Claim Elements*

The district court invalidated the patents on the theory that “what the claim adds to the prior art is the use of rules, rather than artists, to set morph weights and transitions between phonemes.” A18. It stated that even “if . . . the inventive step is the use of timing rules, . . . that still leaves an abstract idea at the point of novelty” because “the user, not the patent” provides the specific rules for any given character. A18-19.

For the reasons given above, the district court erred by stripping away all of the other claim elements. But even so, the claims do not cover merely the "idea" of using rules. A critical component of the claimed method is the use of—and the recognition that it is necessary to use only—very specific ***types*** of rules: those "that define output morph weight set stream ***as a function of* [1] *phoneme sequence and* [2] *time of said phoneme sequence***." A37, 11:30-32 (emphasis added). That was ***not*** in the prior art, and it was highly inventive: Of the "infinitely many" "'sets of rules'" that, according to defendants, "could be developed," D.Ct.Dkt.#338, at 13, Rosenfeld identified the variables that were sufficient to produce high-quality animation automatically. That is a specific "application" of the larger "idea" of using rules in lip-synchronization animation. *Alice*, 134 S. Ct. at 2357. Anyone may use any different rules without fear of infringement. While the district court stated that the patents "preempt the ***field***" of "using a rules-based morph target approach" to "automatic lip synchronization for computer-generated 3-D animation," A19, that ignores the claims' actual, far narrower scope.

The sole reason the district court gave for finding the claimed use of rules to be an "abstract idea," moreover, is that the "concepts are specified at the highest level of generality," and "the user, not the patent," provides the individual rules. A18. But the rules are ***not*** claimed at the highest level of generality. The ***specific types of rules*** are identified, and instructions on how to generate them are pro-

vided. That the ultimate rules are not set forth is a result of the fact that implementation of the patented process necessarily varies from character to character.

Moreover, as discussed above (at 46-47), that concern about "generality" is more appropriately addressed under § 112's "written description" requirement, *see AbbVie*, 759 F.3d at 1299, "wholly apart from whether the invention falls into a category of statutory subject matter" under § 101, *Diehr*, 450 U.S. at 190 (quotation marks omitted). The district court acknowledged that the written-description requirement guards against the very concerns it raised. *See* A10. Yet it never explained why those issues should be addressed under § 101 instead. The claims are not written so broadly as to risk "t[ying] up" any "basic tools of scientific and technological work." *Mayo*, 132 S. Ct. at 1301. The "abstract ideas" exception to § 101 is not implicated.

C. The District Court's Analysis Threatens Software-Based Patents

Software technology provides the "infrastructure" of innovation in today's "information age." President's Info. Tech. Advisory Comm., Nat'l Coordination Office for Computing, Info. & Commc'ns, *Information Technology Research: Investing in Our Future* 23 (1999), *available at* http://research.microsoft.com/en-us/um/people/gray/papers/pitac_report_99_2_24.pdf. Many technological advances that once would have been achieved by altering a device's mechanical structure are now achieved instead through improving the software that controls the device.

Since *Alice*, this Court has regularly invalidated mere business methods implemented using only generic computer functions. *See Ultramercial*, 772 F.3d at 715; *buySAFE*, 765 F.3d at 1355; *Planet Bingo*, 576 F. App'x at 1009. The decision below, however, goes well beyond that. As Judge Pfaelzer explained in *Caltech*, "it is difficult to imagine any software patent that survives" under the approach to § 101 adopted in this case. 2014 WL 5661290, at *11. "[M]ost inventions today build on what is known in the art, and an improvement to software will almost inevitably be an algorithm or concept which, when viewed in isolation, will seem abstract. This analysis would likely render all software patents ineligible . . . ." *Id.*

Consider the example of a patent covering software that makes an anti-lock brake system ("ABS") work better. From any practical perspective, it is hard to see how that could be a mere abstract idea—an ABS system includes such tangible components as the vehicle's physical brakes, wheel speed sensors, an electronic control unit, and a hydraulic modulator unit with pump and valves. But it is software that monitors the rate of rotation of each wheel and determines when the computer should apply the brakes to prevent wheel lock and ensure the driver maintains control of the car. AAA Foundation for Traffic Safety, *FAQs: Anti-Lock Braking System (ABS)* (Feb. 15, 2005), https://www.aaafoundation.org/faqs-anti-lock-braking-system-abs. A better ABS system may mean better software to pro-

cess the informational inputs more quickly and efficiently, and to better decide when and how the computer should apply the brakes.

Under the district court's analysis, an improved ABS system would not be patent-eligible under § 101. According to the district court, "where a claim recites tangible steps, but the only new part of the claim is an abstract idea, that may constitute a claim to an abstract idea." A15. But the "point of novelty" in ABS software will "almost inevitably be an algorithm" or mathematical formula that, "when viewed in isolation, will seem abstract." *Caltech*, 2014 WL 5661290, at *11. There may be nothing tangible about the improved algorithm standing alone. The fact that the software is utilized to control tangible components—brakes, cylinders, pumps, and valves—that moor the claim to a mechanical system, would be deemed irrelevant merely because they are in the prior art. That makes no sense. Where an innovation in software represents a "new and useful advance[] in technology," the patent system should encourage and protect it. *Pfaff*, 525 U.S. at 63. The district court's approach loses sight of the overall purposes of the patent system, and the concerns motivating the abstract-ideas exception.

Like many of today's useful arts, improvements in computer animation are driven by improvements in computer software. The district court's faulty approach jeopardizes the software industry in general, and advances in animation software in

particular. Rosenfeld's invention here, which dramatically improved video animation, is patent-eligible subject matter under § 101. The Court should so hold.

CONCLUSION

The district court's judgment should be reversed.

February 27, 2015

Respectfully submitted,

/s/ Jeffrey A. Lamken

Jeffrey A. Lamken
Counsel of Record
Michael G. Pattillo, Jr.
MOLO LAMKEN LLP
The Watergate, Suite 660
600 New Hampshire Avenue, N.W.
Washington, D.C. 20037
(202) 556-2000 (telephone)
(202) 556-2001 (fax)
jlamken@mololamken.com

John Whealan
4613 Merivale Road
Chevy Chase, MD 20815
(202) 994-2195 (telephone)

Mark S. Raskin
Robert A. Whitman
John F. Petrsoric
MISHCON DE REYA NEW YORK LLP
750 Seventh Avenue 26th Floor
New York, NY 10019
(212) 612-3279 (telephone)
(212) 612-3297 (fax)
mark.raskin@mishcon.com

Counsel for Plaintiff-Appellant McRO, Inc.

ADDENDUM

JS - 6

UNITED STATES DISTRICT COURT
CENTRAL DISTRICT OF CALIFORNIA

McRO, Inc., d.b.a. Planet Blue, Plaintiffs, v. **Namco Bandai Games America, Inc.**, Defendants.	No. CV 12-10322-GW(FFMx) **RULING ON DEFENDANTS' MOTION FOR JUDGMENT ON THE PLEADINGS BASED ON UNPATENTABILITY UNDER 35 U.S.C. § 101**

I. Background

The Court is presiding over two sets of consolidated patent infringement cases filed by Plaintiff McRO, Inc., d.b.a. Planet Blue ("Plaintiff" or "Planet Blue"): the "Track 1" cases, consolidated under Case No. CV-12-10322,[1] and the "Track 2" cases, consolidated under Case No. CV 13-1872.[2] The cases all involve Plaintiff's

[1] The current Track 1 cases are: *McRO, Inc. v. Namco Bandai Games America, Inc.*, CV-12-10322; *McRO, Inc. v. Konami Digital Entertainment, Inc.*, CV-12-10323; *McRO, Inc. v. Sega of America, Inc.*, CV-12-10327; *McRO, Inc. v. Electronics Arts, Inc.*, CV-12-10329; *McRO, Inc. v. Obsidian Entertainment, Inc.*, CV-12-10331; *McRO, Inc. v. Disney Interactive Studios, Inc.*, CV-12-10333; *McRO, Inc. v. Naughty Dog, Inc.*, CV-12-10335; *McRO, Inc. v. Capcom USA, Inc.*, CV-12-10337; *McRO, Inc. v. Square Enix, Inc.*, CV-12-10338; *McRO, Inc. v. Neversoft Entertainment, Inc.*, CV-12-10341; *McRO, Inc. v. Treyarch Corporation*, CV-12-10342; *McRO, Inc. v. Atlus U.S.A., et al.*, CV-13-1870; McRO, *Inc. v. Sucker Punch Productions, LLC*, CV-14-0332; *McRO, Inc. v. Activision Blizzard, Inc.*, CV-14-0336; *McRO, Inc. v. Infinity Ward, Inc.*, CV-14-0352; *McRO, Inc. v. LucasArts Entertainment Company LLC*, CV-14-358; *McRO, Inc. v. Sony Computer Entertainment America, LLC, et al.*, CV-14-0383; *McRO, Inc. v. Warner Bros. Interactive Entertainment Inc.*, CV-14-0417.

[2] The current Track 2 cases are: *McRO, Inc. v. Valve Corporation*, CV-13-1874; *McRO, Inc. v. Codemasters USA Group, Inc. et al*, CV-14-0389; *McRO, Inc. v. Codemasters, Inc., et al*, CV-14-0439.

allegation that Defendants directly or indirectly infringed two patents for automatically animating the lip synchronization and facial expressions of 3D characters. The cases are proceeding on different tracks due to the filing or transfer dates of the cases, although various later-filed cases have been consolidated into Track 1 due to corporate or counsel relationships.

This Motion for Judgment on the Pleadings Based on Unpatentability under 35 U.S.C. § 101 ("Motion") was jointly filed by all defendants in both Tracks: Namco Bandai Games America, Inc.; Sega of America, Inc.; Electronic Arts, Inc.; Disney Interactive Studios, Inc.; Capcom USA, Inc.; Neversoft Entertainment, Inc.; Treyarch Corporation; Warner Bros. Interactive Entertainment, Inc.; LucasArts Entertainment Co. LLC; Activision Publishing, Inc.; Blizzard Entertainment, Inc.; Infinity Ward, Inc.; Atlus U.S.A., Inc.; Konami Digital Entertainment, Inc.; Square Enix, Inc.; Obsidian Entertainment, Inc.; Naughty Dog, Inc.; Sony Computer Entertainment America, LLC; Sucker Punch Productions, LLC; The Codemasters Software Company Limited; Codemasters, Inc.; Codemasters USA Group, Inc.; and Valve Corp. (collectively, "Defendants"). Notice of Mot., Docket No. 338 at 2. Plaintiff filed its Opposition on July 24, 2014. Docket No. 344. Defendants filed their Reply on July 31, 2014. Docket No. 350.

At issue are United States Patent Nos. 6,307,576 ("'576 Patent"), issued October 23, 2001, and 6,611,278 ("'278 Patent"), issued August 26, 2003, both to Maury Rosenfeld, and both titled "Method for Automatically Animating Lip Synchronization and Facial Expression of Animated Characters." The '278 Patent resulted from a continuation of the application that resulted in the '576 Patent, meaning the patents share the same disclosure. *See PowerOasis, Inc. v. T-Mobile USA, Inc.*, 522 F.3d 1299, 1304, n.3 (Fed. Cir. 2008).

The patents explain that prior methods of animating lip synchronization and facial expressions were laborious and uneconomical. '576 Patent 1:14-31. The

patents address that problem with an automated method of using "weighted morph targets and time aligned phonetic transcriptions of recorded text, and other time aligned data." '576 Patent 2:64-3:12. The patents explain that in the relevant art, "'phonemes [are] defined as the smallest unit of speech, and correspond[] to a single sound." '576 Patent 1:34-36. A sound recording can be transcribed into a "time aligned phonetic transcription" in which the timing of each phoneme is noted. '576 Patent 1:32-34. Such transcriptions can be created by hand or by automatic speech recognition programs. '576 Patent 1:39-43.

The patents explain that the prior art practice for 3-D computer generated speech animation was by manual techniques using a "morph target" approach. '576 Patent 1:44-46. That approach uses a reference model of a neutral mouth position in conjunction with "morph targets," which are models of the mouth in non-neutral positions corresponding to different phonemes. '576 Patent 1:46-49. The reference model and morph targets all share the same "topology" of the mouth, defined by the same number and placement of "vertices" that designate specific points on the mouth. For example, vertex "n" on the neutral mouth and all of the morph targets may represent the left corner of the mouth. '576 Patent 1:51-54.

The "deltas," or changes, of each vertex on each morph target relative to the corresponding vertex on the neutral model are computed as a vector to produce an individual "delta set" of vectors for each morph target. '576 Patent 1:58-62. From the neutral model, the animator need not move the mouth position all the way to a morph target. Instead, the animator can apply a value between 0 and 1, called the "morph weight," to a delta set to move the mouth just a percentage of the way to the corresponding morph target. '576 Patent 1:63-2:1. For example, if the sound (morph target) is "oh," and the morph weight is 0.5, the mouth only moves halfway between the neutral position and the "oh" morph target. '576 Patent 2:16-22. It is also possible to blend the morph targets, for example, 0.3 "oh" and 0.7 "ee," resulting in

a mouth position exhibiting a combination of the "oh" and "ee" sound characteristics. '576 Patent 2:23-28.

According to the patents, applying the appropriate morph weights in the prior art was usually done using a "keyframe" approach. In the keyframe approach, an artist sets the morph weights at certain important times, and a computer program then interpolates each of the channels at each frame between the keyframes. '576 Patent 2:29-34. The patents state that this method requires the artist to manually set a large number of keyframes, which is tedious, time consuming, and inaccurate. '576 Patent 2:34-37. Therefore, an object of the invention is to provide "an extremely rapid and cost effective means to automatically create lip synchronization and facial expression in three dimensional animated characters." '576 Patent 2:50-54.

The invention "utilizes a set of rules that determine the system[']s output comprising a stream or streams of morph weight sets when a sequence of timed phonemes or other timed data is encountered." '576 Patent 3:3-7. The invention includes:

> [C]onfiguring a set of default correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and specifying a plurality of morph weight set transition rules for specifying durational data for the generation of transitionary curves between the plurality of morph weight sets, allowing for the production of a stream of specified morph weight sets to be processed by a computer animation system . . . .

'576 Patent 3:23-30.

Defendants argue that the claims of both patents in suit are patent ineligible under 35 U.S.C. § 101 because they merely "set[] forth the previously-known animation method as a series of mathematical steps, and instruct[] the user to perform those steps on a computer." Mot., Docket No. 338 at 12.

II. Legal Standard

A. Motion for Judgment on the Pleadings

Rule 12(c) of the Federal Rules of Civil Procedure permits a party to move to dismiss a suit "[a]fter the pleadings are closed . . . but early enough not to delay trial."

Fed. R. Civ. P. 12(c). "Judgment on the pleadings is proper when, taking all allegations in the pleading as true, the moving party is entitled to judgment as a matter of law." *Stanley v. Trustees of Cal. State Univ.*, 433 F.3d 1129, 1133 (9th Cir. 2006); *see also Fleming v. Pickard*, 581 F.3d 922, 925 (9th Cir. 2009). Because a motion for judgment on the pleadings is "functionally identical" to a motion to dismiss, the standard for a Rule 12(c) motion is the same as for a Rule 12(b)(6) motion. *See Platt Elec. Supply, Inc. v. EOFF Elec., Inc.*, 522 F.3d 1049, 1052 n.1 (9th Cir. 2008).

A complaint may be dismissed for failure to state a claim upon which relief can be granted for one of two reasons: (1) lack of a cognizable legal theory or (2) insufficient facts under a cognizable legal theory. *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 555 (2007). *See also Mendiondo v. Centinela Hosp. Med. Ctr.,* 521 F.3d 1097, 1104 (9th Cir. 2008) ("Dismissal under Rule 12(b)(6) is appropriate only where the complaint lacks a cognizable legal theory or sufficient facts to support a cognizable legal theory."). A motion to dismiss should be granted if the complaint does not proffer enough facts to state a claim for relief that is plausible on its face. *See Twombly*, 550 U.S. at 558-59, 570; *see also William O. Gilley Enters., Inc. v. Atlantic Richfield Co.*, 588 F.3d 659, 667 (9th Cir. 2009) (confirming that *Twombly* pleading requirements "apply in all civil cases"). "[W]here the well-pleaded facts do not permit the court to infer more than the mere possibility of misconduct, the complaint has alleged – but it has not 'show[n]' – 'that the pleader is entitled to relief.'" *Ashcroft v. Iqbal*, 556 U.S. 662, 679 (2009) (quoting Fed. R. Civ. P. 8(a)(2)).

In deciding a 12(b)(6) or 12(c) motion, the court is limited to the allegations on the face of the complaint (including documents attached thereto), matters which are properly judicially noticeable and other extrinsic documents when "the plaintiff's claim depends on the contents of a document, the defendant attaches the document to its motion to dismiss, and the parties do not dispute the authenticity of the document, even though the plaintiff does not explicitly allege the contents of that

document in the complaint." *Knievel v. ESPN*, 393 F.3d 1068, 1076 (9th Cir. 2005). The court must construe the complaint in the light most favorable to the plaintiff and must accept all factual allegations as true. *Cahill v. Liberty Mutual Ins. Co.*, 80 F.3d 336, 337-38 (9th Cir. 1996). The court must also accept as true all reasonable inferences to be drawn from the material allegations in the complaint. *See Brown v. Elec. Arts, Inc.*, 724 F.3d 1235, 1247-48 (9th Cir. 2013); *Pareto v. F.D.I.C.*, 139 F.3d 696, 699 (9th Cir. 1998). Conclusory statements, unlike proper factual allegations, are not entitled to a presumption of truth. *See Iqbal*, 556 U.S. at 681; *Moss v. U.S. Secret Serv.*, 572 F.3d 962, 969 (9th Cir. 2009).

B. Patentable Subject Matter Under 35 U.S.C. § 101[3]

35 U.S.C. § 101 "defines the subject matter that may be patented under the Patent Act." *Bilski v. Kappos*, 561 U.S. 593, ___, 130 S.Ct. 3218, 3225 (2010). It provides:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Id. "In choosing such expansive terms . . . modified by the comprehensive 'any,' Congress plainly contemplated that the patent laws would be given wide scope" "to ensure that 'ingenuity should receive a liberal encouragement.'" *Id.* (quoting *Diamond v. Chakrabarty*, 447 U.S. 303, 308 (quoting 5 Writings of Thomas Jefferson 75–76 (H. Washington ed. 1871)) (some internal quotation marks omitted).

The "wide scope" of patent eligibility is not unlimited. Instead, the Supreme Court has invented or discovered "three specific exceptions to § 101's broad patent-eligibility principles: 'laws of nature, physical phenomena, and abstract ideas.'" *Bilski*, 130 S.Ct. at 3225 (quoting *Chakrabarty*, 447 U.S. at 309). Although "the exceptions have defined the statute's reach as a matter of statutory *stare decisis*

[3] This section concerning the applicable legal standard is the same as the corresponding section in this Court's recent decision in *Eclipse IP LLC v. McKinley Equip. Corp.*, CV-14-154-GW (AJWx), 2014 WL 4407592 (C.D. Cal. Sept. 4, 2014), except for minor changes.

going back 150 years,"[4] *id*., they have not been enumerated consistently during that time. Forty years ago, the list of unpatentable "basic tools of scientific and technological work" was: "[p]henomena of nature . . . , mental processes, and abstract intellectual concepts." *Gottschalk v. Benson*, 409 U.S. 63, 67 (1972).

In *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*, 132 S.Ct. 1289 (2012), the Supreme Court "set forth a framework for distinguishing patents that claim laws of nature, natural phenomena, and abstract ideas from those that claim patent-eligible applications of those concepts." *Alice Corp. Pty. Ltd. v. CLS Bank Int'l*, 134 S. Ct. 2347, 2355 (2014). That framework is as follows:

> First, we determine whether the claims at issue are directed to one of those patent-ineligible concepts. If so, we then ask, "[w]hat else is there in the claims before us?" To answer that question, we consider the elements of each claim both individually and "as an ordered combination" to determine whether the additional elements "transform the nature of the claim" into a patent-eligible application. We have described step two of this analysis as a search for an "'inventive concept'" – *i.e.*, an element or combination of elements that is "sufficient to ensure that the patent in practice amounts to significantly more than a patent upon the [ineligible concept] itself."

Id. at 2355 (citations omitted).

Describing this as a two-step test may overstate the number of steps involved. If the claim is not "directed" to a patent-ineligible concept, then the test stops at step one. If the claim is so directed, but we find in step two that the claim contains an "inventive concept" that "transforms" the nature of the claim into something patent eligible, then it seems that there was a categorization error in finding the claim – which is considered "as an ordered combination" – "directed to an abstract idea" in step one.

---

[4] "Statutory *stare decisis*" is a recent coinage, apparently used for the first time by Justice Scalia concurring in part in *Rita v. United States*, 551 U.S. 338, 368 (2007). Justice Ginsburg was the next to use the phrase: "Although I joined Justice SCALIA in *Rita* accepting the *Booker* remedial opinion as a matter of 'statutory *stare decisis*' . . . ." *Kimbrough v. United States*, 552 U.S. 85, 116 (2007). Justice Ginsburg's use of quotation marks could have been a comment on the novelty of the phrase, but might have simply indicated a quotation. In any event, Justice Ginsburg later used the phrase without quotation marks in *CSX Transp., Inc. v. McBride*, 131 S. Ct. 2630, 2641 (2011). The context there makes clear that the phrase refers to the principle that "[c]onsiderations of *stare decisis* have special force in the area of statutory interpretation, for here, unlike in the context of constitutional interpretation, the legislative power is implicated, and Congress remains free to alter what we have done." *Patterson v. McLean Credit Union*, 491 U.S. 164, 172-73 (1989).

So, the two-step test may be more like a one step test evocative of Justice Stewart's most famous phrase. *See Jacobellis v. State of Ohio*, 378 U.S. 184, 197 (1964) (Stewart, J. concurring) ("I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it . . . ."); *cf. Alice*, 134 S.Ct. at 2357 ("In any event, we need not labor to delimit the precise contours of the 'abstract ideas' category in this case.").

Rest and relaxation prevailed in *Alice* because it was "enough to recognize that there is no meaningful distinction between the concept of risk hedging in *Bilski* and the concept of intermediated settlement at issue [in *Alice*]. Both are squarely within the realm of 'abstract ideas' . . . ." *Id*. at 2357 (citing to *Bilski*, 130 S.Ct. 3218). Thus, so far, the two-part test for identifying an abstract idea appears to be of limited utility, while comparisons to previously adjudicated patents – or more precisely, to past cases' characterizations of those patents[5] – have done the heavy lifting. *See also Bilski*, 130 S. Ct. at 3229 ("Rather than adopting categorical rules that might have wide-ranging and unforeseen impacts, the Court resolves this case narrowly on the basis of this Court's decisions in *Benson*, *Flook*, and *Diehr* . . . .").[6] It remains true that "[t]he life of the law has not been logic: it has been experience." Oliver Wendell Holmes, Jr., *The Common Law* 1 (1881).

But despite its narrow holding, *Alice* did categorically establish a clear rule that had previously been subject to debate: "mere recitation of a generic computer cannot transform a patent-ineligible abstract idea into a patent-eligible invention." 134 S.Ct. at 2358. And before *Alice*, it was unclear to some, including the USPTO,

---

[5] *Mayo* noted that, as to the patent-ineligible approach of simply instructing artisans "to apply" unpatentable subject matter, "[t]he process in *Diehr* was not so **characterized**; that in *Flook* was **characterized** in roughly this way." 132 S. Ct. at 1299-1300 (emphasis added).

[6] Scholars have argued that "the *Mayo* decision has revived the *Flook* approach, although without displacing *Diehr* or explaining how the two apparently contradictory decisions can be reconciled." Brief of Professors Peter S. Menell and Jeffrey A. Lefstin as Amici Curiae in Support of Respondents, *Alice Corp. Pty, Ltd. v. CLS Bank Int'l*, No. 13-298, 2014 U.S. Briefs LEXIS 784 at 10 (Feb. 27, 2014).

that the framework set forth in *Mayo* applied to abstract ideas as well as to the law of nature/natural phenomena at issue in *Mayo*. *See* Memo to Patent Examining Corps from Andrew H. Hirschfeld, Deputy Commissioner for Patent Examination Policy, Preliminary Examination Instructions in view of the Supreme Court Decision in *Alice Corporation Pty. Ltd. v. CLS Bank International, et al.* (June 25, 2014), *available at* http://www.uspto.gov/patents/announce/alice_pec_25jun2014.pdf.[7]

And, while the boundaries of the judicial exceptions remain subject to further development, the Supreme Court has clearly stated the policy underlying those exceptions, i.e. avoiding patents that "too broadly preempt the use of a natural law [or abstract idea]." *Mayo*, 132 S.Ct. at 1294. Thus, patent law should "not inhibit further discovery by improperly tying up the future use of laws of nature [or abstract ideas]." *Id*. at 1301.

Mayo discussed the Supreme Court's 1854 decision upholding many of Samuel Morse's telegraph patent claims, but invalidating the most general claim, which covered "the use of the motive power of the electric or galvanic current . . . however developed, for making or printing intelligible characters, letters, or signs, at any distances." *Id*. The Supreme Court presciently explained that such a claim would inhibit, rather than promote, the progress of the useful arts:

> For aught that we now know some future inventor, in the onward march of science, may discover a mode of writing or printing at a distance by means of the electric or galvanic current, without using any part of the process or combination set forth in the plaintiff's specification. His invention may be less complicated – less liable to get out of order – less expensive in construction, and in its operation. But yet if it is covered by this patent the inventor could not use it, nor the public have the benefit of it without the permission of this patentee.

Id. (quoting *O'Reilly v. Morse*, 15 How. 62, 113 (1854).) True, patents always

---

[7] Indeed, in the USPTO's view, *Alice*'s embrace of the *Mayo* framework for abstract idea cases was such a significant change or clarification that it has withdrawn issued notices of allowance – that is, stopped patents that had made it all the way through examination and were about to issue – "due to the presence of at least one claim having an abstract idea and no more than a generic computer to perform generic computer functions." USPTO Commissioner for Patents Peggy Focarino, Update on USPTO's Implementation of 'Alice v. CLS Bank' (Aug. 4, 2014), available at http://www.uspto.gov/blog/director/entry/update_on_uspto_s_implementation.

present some impediment to follow-on innovation. The principle is one of balance: patents should not "foreclose[] more future invention than the underlying discovery could reasonably justify." *Mayo*, 132 S.Ct. at 1301.

Of course, § 101 is not the sole, or even primary, tool to ensure that balance. Every condition of patentability set forth in the Patent Act acts to ensure that patents promote, rather than retard, the progress of science and useful arts. For example, in a manner quite similar to recent § 101 jurisprudence, "[t]he written description requirement guards against claims that 'merely recite a description of the problem to be solved while claiming all solutions to it and . . . cover any compound later actually invented and determined to fall within the claim's functional boundaries.'" *Abbvie Deutschland GmbH & Co., KG v. Janssen Biotech, Inc.*, __ F.3d __, 2013-1338, 2014 WL 2937477, 11 (Fed. Cir. July 1, 2014) (quoting *Ariad Pharm., Inc. v. Eli Lilly & Co.*, 598 F.3d 1336, 1353 (Fed. Cir. 2010)).

However, scholars have argued that the written description and enablement doctrines of § 112, as currently applied, do not adequately prevent unwarranted obstructions to follow-on innovation, and have urged that § 101 can and should do so. *See, e.g.*, Lemley et al., *Life After Bilski*, 63 Stan. L. Rev. 1315, 1330 (2011) (cited in *Mayo*, 132 S.Ct. at 1301-03, 1304); *but see* Lemley, *Point of Novelty*, 105 Nw. U. L. Rev. 1253, 1279 (2011) ("[T]here is good reason to worry about overbroad patent claims that lock up a wide swath of potential future applications. But the enablement and written description doctrines largely address that concern.").

In any event, the Supreme Court has spoken, and § 101 now plays an important limiting role. But District Courts and the Federal Circuit are now left with the task of figuring out when the "two-part" test is satisfied. Perhaps something like the function-way-result test used to evaluate infringement under the doctrine of equivalents might be useful. Thus, in one long-standing formulation, an accused instrumentality infringes "if it performs substantially the same function in

substantially the same way to obtain the same result." *Union Paper-Bag Mach. Co. v. Murphy*, 97 U.S. 120, 125 (1877); *InTouch Technologies, Inc. v. VGO Commc'ns, Inc.*, 751 F.3d 1327, 1343 (Fed. Cir. 2014).

The test in practice often focuses on the "way" aspect of the test, because function and result are often identical in the patent and accused product, and the question is whether the accused infringer uses the same "way." Laura A. Handley, *Refining the Graver Tank Analysis with Hypothetical Claims: A Biotechnology Exemplar*, 5 Harv. J.L. & Tech. 36 (1991) ("In practice, the second prong of the test – 'substantially the same way' is often emphasized, since most infringement suits result from competition for a given market niche which dictates the 'function' and 'result' prongs.") (citing *Perkin-Elmer Corp. v. Westinghouse Elec. Corp.*, 822 F.2d 1528, 1531 (Fed. Cir. 1987)).[8]

Similarly, the question in the abstract idea context is whether there are other *ways* to use the abstract idea in the same field. If so, the Supreme Court has expressly encouraged others to find those other ways, without being held back by patents that preempt the whole concept. *Mayo*, 132 S.Ct. at 1294 (citing *O'Reilly*, 15 How. at 113); *Alice*, 134 S.Ct. at 3258 (noting "the pre-emption concern that undergirds our § 101 jurisprudence.").

Concomitantly, we must be wary of facile arguments that a patent preempts all applications of an idea. It may often be easier for an infringer to argue that a patent fails § 101 than to figure out a different way to implement an idea, especially a way that is "less complicated – less liable to get out of order – less expensive in construction, and in its operation." *O'Reilly*, 15 How. at 113. But the patent law does not privilege the leisure of an infringer over the labors of an inventor. Patents

[8] *Perkin-Elmer* held that "repeated assertions that the claimed and accused devices perform substantially the same function and achieve substantially the same end result are not helpful. That circumstance is commonplace when the devices are sold in competition. That a claimed invention and an accused device may perform substantially the same function and may achieve the same result will not make the latter an infringement under the doctrine of equivalents where it performs the function and achieves the result in a substantially different way." 822 F.2d at 1532 n.6.

should not be casually discarded as failing § 101 just because the infringer would prefer to avoid the work required to develop non-infringing uses of the abstract idea at the heart of an appropriately circumscribed invention.

III. <u>Analysis</u>

A. Defendants' Patents Are Irrelevant

Plaintiff argues that Defendants' own patents for lip-synchronization, some of which issued very recently, undermine Defendants' argument that the patents-in-suit are directed to unpatentable subject matter. Opp'n, Docket No. 344 at 20-22. The validity of Defendants' patents is not before the Court, and Plaintiff has cited no authority for the proposition that Defendants' obtaining them operates as an estoppel in this case. There may be numerous factual differences between Defendants' patents and those at issue here. And even if Defendants' patents rise and fall with Plaintiff's, it is hard to fault anyone for seeking patents that may turn out to be invalid where the applicable standards are shifting and uncertain. "A change in the weather has known to be extreme." Bob Dylan, *You're a Big Girl Now*, <u>Blood on the Tracks</u> (Columbia Records 1974).

B. The Patents-in-Suit Fail § 101

1. The Claims, In Isolation, Appear Tangible and Specific

Defendants argue that the patents-in-suit are directed to a "fundamental, abstract animation practice," namely, "the abstract idea of rules-based synchronization of animated mouth movement." Mot., Docket No. 338 at 12. That is, Defendants argue that the patents cover the mere idea of using rules for three-dimensional lip synchronization, without requiring specific content for those rules. *Id*. at 12-13. But considered standing alone, the asserted claims do not seem to cover any and all use of rules for three-dimensional lip synchronization. The independent claims of each of the patents in suit are:

'576 Patent claim 1:

> A method for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:
> obtaining a first set of rules that define output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence;
> obtaining a timed data file of phonemes having a plurality of sub-sequences;
> generating an intermediate stream of output morph weight sets and a plurality of transition parameters between two adjacent morph weight sets by evaluating said plurality of sub-sequences against said first set of rules;
> generating a final stream of output morph weight sets at a desired frame rate from said intermediate stream of output morph weight sets and said plurality of transition parameters; and
> applying said final stream of output morph weight sets to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters.

'278 Patent claim 1:

> A method for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:
> obtaining a first set of rules that defines a morph weight set stream as a function of phoneme sequence and times associated with said phoneme sequence;
> obtaining a plurality of sub-sequences of timed phonemes corresponding to a desired audio sequence for said three-dimensional characters;
> generating an output morph weight set stream by applying said first set of rules to each sub-sequence of said plurality of sub-sequences of timed phonemes; and
> applying said output morph weight set stream to an input sequence of animated characters to generate an output sequence of animated characters with lip and facial expression synchronized to said audio sequence.

Facially, these claims do not seem directed to an abstract idea. They are tangible, each covering an approach to automated three-dimensional computer animation, which is a specific technological process. They do not claim a monopoly, as Defendants argue, on "the idea that the human mouth looks a certain way while speaking particular sounds," "applied to the field of animation." Mot., Docket No. 338 at 12, n.9. Further, the patents do not cover the prior art methods of computer assisted, but non-automated, lip synchronization for three-dimensional computer animation.

And according to Defendants, they do not cover the automated methods of lip synchronization for three-dimensional computer animation that Defendants employ. It is hard to show that an abstract idea has been preempted if there are noninfringing ways to use it in the same field. Section 101 motions can place parties in unfamiliar and uncomfortable positions: here it is to the patentee's advantage to identify noninfringing alternatives, and it is the accused infringer's advantage to posit the lack of any; the reverse of their positions at the infringement and damages stages of the case.

At first blush, it is therefore difficult to see how the claims might implicate the "basic underlying concern that these patents tie up too much future use of" any abstract idea they apply. *Mayo*, 132 S. Ct. at 1302; *Alice*, 134 S.Ct. at 2358 (noting "the pre-emption concern that undergirds our § 101 jurisprudence").

2. The Claims Must Be Evaluated in the Context of the Prior Art

However, for purposes of the § 101 analysis, it is not enough to view the claims in isolation. Instead, when determining whether a patent contains an adequate inventive concept, the Court must factor out conventional activity. That is because the inclusion of "well-understood, routine, conventional activity" previously used in the field "is normally not sufficient to transform an unpatentable law of nature [or abstract idea] into a patent-eligible application . . . ." *Mayo*, 132 S.Ct. at 1298.[9] Further, in addition to evaluating each step of the claim, the claims must be considered as "an ordered combination." *Alice*, 132 S.Ct. at 2355.

This dual analysis tracks the law's long-standing concern with patents that

[9] In a forthcoming paper, Jeffrey Lefstin argues that for more than a hundred years, the lesson drawn from the English *Neilson* case (relied upon by the Supreme Court in *Mayo*) was that any practical application of a new discovery was patentable, even if the application was entirely conventional. Jeffrey Lefstin, *Inventive Application: A History*, Fla. L. Rev. & Hastings Research, Paper No. 94 (Mar. 2014), *available at* http://ssrn.com/abstract =2398696. This is contrary to the current law that "appending conventional steps, specified at a high level of generality, to laws of nature, natural phenomena, and abstract ideas cannot make those laws, phenomena, and ideas patentable." *Mayo*, 132 S.Ct. at 1300. What the Supreme Court says about prior cases is often more important than what the cases themselves said. *See, e.g., Daimler AG v. Bauman*, 134 S. Ct. 746, 756 n.8 (2014) (eight-member majority chiding Justice Sotomayor for relying in her concurrence on the facts recited in *Perkins v. Benguet Consol. Mining Co.*, 342 U.S. 437 (1952) and in the intermediate appellate opinion in that case, rather than acquiescing to the characterization of *Perkins* in a recent decision, *Goodyear Dunlop Tires Operations, S.A. v. Brown*, 131 S.Ct. 2846 (2011)) (which Justice Sotomayor had joined).)

consist of old material with the addition of a new, but abstract, idea: "the vice of a functional claim exists not only when a claim is 'wholly' functional, if that is ever true, but also when the inventor is painstaking when he recites what has already been seen, and then uses conveniently functional language at the exact point of novelty." *Gen. Elec. Co. v. Wabash Appliance Corp.*, 304 U.S. 364, 371, 58 S. Ct. 899, 903 (1938). An abstract idea is the extreme case of functional language.

Thus, where a claim recites tangible steps, but the only new part of the claim is an abstract idea, that may constitute a claim to an abstract idea. *See Alice*, 134 S. Ct. at 2358. (disregarding the presence of a computer in the claim given "the ubiquity of computers"); *Mayo*, 132 S.Ct. at 1297-98 (claim step calling for administration of drug only disregarded because it "refers to the relevant audience, namely doctors who treat patients with certain diseases with thiopurine drugs"; claim step of determining the level of the relevant metabolites disregarded because it was "well known in the art").

Here, the patents teach that in the prior art, three-dimensional character lip synchronization was performed using a "timed data file of phonemes having a plurality of sub-sequences," as recited in the claims. '576 Patent 1:32-43. But the prior art did not, according to the patents, involve obtaining rules that define output morph weight sets as a function of the phonemes, or using those rules to generate the morph weight sets. Instead, an artist manually set the morph weights at certain important keyframes, and a computer program then interpolated the frames between the keyframes. '576 Patent 2:29-37. Therefore, while tangible, the steps of (1) using a timed phoneme transcript, (2) setting morph weight sets at keyframes, or (3) interpolating between keyframes, are not "inventive steps" that could transform the claims herein into patent eligible subject matter, if those claims are directed to an abstract idea.

In attacking the claims as simply drawn to the abstract idea of "rules-based lip-

synchronized animation on a computer," Mot., Docket No. 338 at 3, Defendants' argument does not account for the presence in the claims, or the Court's construction, of "morph weight set." The Court construed "morph weight set" as a "set of values, one for each delta set, that, when applied, transform the neutral model to some desired state, wherein each delta set is the [set of vectors] from each vertex on the neutral (reference) model to each vertex on a model of another mouth position." Rulings on Claim Constr., Docket No. 298-1 at 9.

However, the patents themselves teach that the prior art includes using morph targets that correspond to phonemes and calculating delta sets that contain the vectors from each vertex on the neutral model to the morph target. '576 Patent at 1:44-62. So, while Defendant's characterization is overly broad, it would be fair to characterize the claims as drawn to the idea of automated rules-based use of morph targets and delta sets for lip-synchronized three-dimensional animation. Indeed, Plaintiff's expert opines that:

> A central part of the creative insight of the patents is the realization to use the specific approach of using morph weight set representations of the facial shape coupled with rules, including explicit and distinct timing rules, to generate keyframes. This approach uniquely provides the automation required to produce animation in a cost-effective way, yet provided the necessary artistic control required to produce commercial grade animation.

Declaration of Michael Gleicher, Ph.D. in Supp. of Opp'n, Docket No. 345, ¶ 20. Defendants object to this testimony, because "[t]he Court may not consider declarations in opposition to a Rule 12(c) motion without converting the motion to a motion for summary judgment." Defs.' Objections to Declarations Filed in Connection with Motion for Judgment on the Pleadings, Docket No. 351 at 2.[10] It is unclear how that response helps Defendants. Certainly, one option is for the Court to deny the Motion as presenting an issue that turns on the facts.

However, nothing in the Declaration affects the analysis. In the paragraph

[10] Plaintiff submitted a response to Defendant's Objections, which also included an unauthorized five-page sur-reply, which the Court would not consider. Planet Blue's Response to Defs.' Objections to Declarations Filed in Opposition to Motion for Judgment on the Pleadings, Docket No. 355. Neither would the Court consider Defendants' Reply to that Response, Docket No. 356.

quoted above, Plaintiff's expert opines that a central part of the patents is "using morph weight set representations of the facial shape coupled with rules, including explicit and distinct timing rules, to generate keyframes." Everyone appears to agree with that characterization, except that Defendants point out that no particular "explicit and distinct" rules are required by the claims. The question is therefore whether the inclusion of that *concept* in the claims satisfies § 101 given (1) the prior art, and (2) the fact that the claims do not require any particular rules.

A consideration of the prior art recited in the patents shows that the point of novelty here is the idea of using rules, including timing rules, to automate the process of generating keyframes. The following chart compares the '576 Patent's claim elements to the prior art described in that patent.

'576 Patent, Claim 1	
Step	**Admitted Prior Art**
A method for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:	Automating the process is the focus of the invention. However, the patent teaches that in the prior art, the use of computerized interpolation partially automated the process by allowing animators to set mouth shapes only at keyframes, rather than at every frame, as would be the case in hand-drawn animation. '576 Patent 2:31-34.
obtaining a first set of rules that define output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence;	Rules for defining morph weight sets as a function of phoneme sequence are disclosed as within the prior art. '576 Patent 1:44-2:28. Rules for defining morph weight sets as a function of timing are not; instead, the timing results from the artist's choice of keyframes. '576 Patent 2:29-34. Note, however, that no particular timing rules are required by any claim.

Step	Admitted Prior Art
generating an intermediate stream of output morph weight sets and a plurality of transition parameters between two adjacent morph weight sets by evaluating said plurality of sub-sequences against said first set of rules;	An intermediate stream of morph weight sets is disclosed as being part of the prior art through the keyframes manually set by the artist. '576 Patent 2:29-34. The transition parameters are not. Those parameters flow from the timing rules.
generating a final stream of output morph weight sets at a desired frame rate from said intermediate stream of output morph weight sets and said plurality of transition parameters; and	The patent teaches that the prior art generated the final stream by interpolating between the keyframes. '576 Patent 2:29-34. Again, transition parameters are not disclosed as being within the prior art.
applying said final stream of output morph weight sets to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters.	Both the final set of output morph weight sets and applying those sets are covered by the interpolation process of the prior art. '576 Patent 2:29-34.

So, what the claim adds to the prior art is the use of rules, rather than artists, to set the morph weights and transitions between phonemes. However, both of these concepts are specified at the highest level of generality. At the hearing on the Motion, Plaintiff emphasized that the rules inventively take into account the timing of the phoneme sequence. But the specification states clearly that "[i]n operation and use, the user must manually set up default correspondence rules" that "specify the durational information needed to generate appropriate transitionary curves between morph weight sets, such as transition start and end times." '576 Patent 6:46-54. Thus, the user, not the patent, provides the rules. And while the patent does provide an example of a very partial set of default and secondary rules, it expressly states that "this is only an example of a set of rules which could be use[d] for illustrative purposes, and many other rules could be specified according to the method of the invention." '576 Patent 7:36-9:23. Because the claim purports to cover all such rules, in light of the prior art, the claim merely states "an abstract idea while adding the words 'apply it.'" *Alice*, 134 S. Ct. at 2358 (quoting *Mayo*, 132 S.Ct. at 1294)

(some quotation marks omitted). The same is true for claim 1 of the '278 Patent, which does not differ in a manner relevant to this analysis.

Here, while the patents do not preempt the field of automatic lip synchronization for computer-generated 3D animation, they do preempt the field of such lip synchronization using a rules-based morph target approach. And if, as Plaintiff suggests, the inventive step is the use of timing rules, given the state of the prior art, that still leaves an abstract idea at the point of novelty, and preventing the development of any additional ways to use that abstract idea in the relevant field. *See Alice*, 134 S. Ct. at 2360 ("the claims at issue amount to 'nothing significantly more' than an instruction to apply the abstract idea of intermediated settlement using some unspecified, generic computer").

3. The Failure of the Claims Is Not Inconsistent with the Inventor Having Developed an Innovative Process

Defendants argue that a "patentee simpl[y] does not waste the time, money and effort to prosecute a patent application for an invention they casually indicate was known in the art." Opp'n, Docket No. 344 at 10-11. But a § 101 defect does not mean that the invention was in the prior art. The invention here may have been novel, but the claims are directed to an abstract idea. And the patent's casual – and honest – description of the prior art was made at a time when, under the then-prevalent interpretation of the law, such admissions were unlikely to be harmful. One unintended consequence of *Alice*, and perhaps of this and other decisions to come, is an incentive for patent applicants to say as little as possible about the prior art in their applications.[11]

Plaintiff points to one Defendant's contemporaneous characterization of Plaintiff's system as "revolutionary." Opp'n, Docket No. 344 at 1 (quoting Decl. of John Petrsoric In Opp'n to Mot., Docket No. 346, Ex. 2, January 27, 1999 Warner

[11]However, that strategy is limited by the doctrine of inequitable conduct.

Bros. Memorandum (inviting colleagues to a demonstration of Plaintiff's "revolutionary lip synch technique" that "utilizes proprietary software.")).

This argument is unpersuasive in this context for two reasons. First, for purposes of the § 101 inquiry, which is different from the § 103 inquiry, the revolutionary nature of an abstract idea does not weigh in favor of patentability. *See Mayo*, 132 S. Ct. at 1293 ("Einstein could not patent his celebrated law that $E=mc^2$ . . . . Such discoveries are 'manifestations of . . . nature, free to all men and reserved exclusively to none.'") (quoting *Chakrabarty*, 100 S.Ct. at 2204). Second, there has been no showing that the cited praise relates to the claims in all their breadth, rather than to a particular implementation that is not specified by the claims. Thus, the inventor's specific implementation of the abstract idea represented by the claim may have been of significant value beyond that of the abstract idea itself.

4. None of the Additional Content in the Asserted Dependent Claims Yields a Different Result

Plaintiff has asserted '576 Patent claims 1, 7-9, and 13, and '278 Patent, claims 1-4, 6, 9, 13, 15-17. Mot., Docket No. 338 at 2. The additional content of the dependent claims is addressed in the following chart:

Claim	Language	Analysis
'576 Patent claim 7	The method of claim 1 wherein said timed data is a time[] aligned phonetic transcriptions data.	Because "time aligned phonetic transcriptions" were used in the prior art ('576 Patent 1:32-37), the additional limitation of this claim does not affect the § 101 analysis.
'576 Patent claim 8	The method of claim 7 wherein said timed data further comprises time aligned data.	This adds nothing to claim 7, and so does not affect the § 101 analysis.

Claim	Language	Analysis
'576 Patent claim 9	The method of claim 7 wherein said timed data further comprises time aligned emotional transcription data.	Not specifically referenced in the patent's description of the prior art. However, this is just another idea of a factor that could be taken into account by the rules; the patent claims no specific method of doing so.
'576 Patent claim 13	The method of claim 1 wherein said first set of rules comprises: correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and morph weight set transition rules specifying durational data for generating transitionary curves between morph weight sets.	Claim 1 already includes "obtaining a first set of rules that define output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence." The specific content of claim 13 is not meaningfully different from that from a § 101 perspective.
'278 Patent claim 2	The method of claim 1, wherein said first set of rules comprises: correspondence rules between all visual phoneme groups and morph weight sets; and morph weight set transition rules specifying durational data between morph weight sets.	These elements have already been discussed in the context of the '576 Patent.
'278 Patent claim 3	The method of claim 2, wherein said durational data comprises transition start and transition end times.	Transition start and end times are inherent in "transition rules specifying durational data between morph weight sets," which is an element of '278 Patent claim 2.
'278 Patent claim 4	The method of claim 1, wherein said desired audio sequence is from a pre-recorded live performance.	This is merely limiting the claim to a particular field of use. "[T]he prohibition against patenting abstract ideas 'cannot be circumvented by attempting to limit the use of the formula to a particular technological environment' . . . ." *Bilski*, 130 S.Ct. at 3230 (quoting *Diehr*, 450 U.S. at 191).
'278 Patent claim 6	The method of claim 1, wherein said plurality of subsequences of timed phonemes is obtained from a file.	This presents the same issue as '278 Patent claim 4. *See* discussion above.

Claim	Language	Analysis
'278 Patent claim 9	The method of claim 1, wherein said generating said output morph weight stream comprises: generating an appropriate morph weight set corresponding to each subsequence of said timed phonemes; and generating time parameters for transition of said appropriate morph weight set from a morph weight set of a prior sub-sequence of said timed data.	This presents the same issue as '278 Patent claim 2. *See* discussion above.
'278 Patent claim 13	The method of claim 1, wherein said plurality of subsequences of timed phonemes comprises a time[] aligned phonetic transcriptions sequence.	This is a basic feature of the prior art. '278 Patent 1:35-47.
'278 Patent claim 15	The method of claim 13, wherein said plurality of subsequences of timed phonemes further comprises time aligned emotional transcription data.	Not specifically referenced in the patent's description of the prior art. However, this is just another idea of a factor that could be taken into account by the rules; the patent claims no specific method of doing so.
'278 Patent claim 16	The method of claim 9, wherein said transition parameters comprises: transition start time; and transition end time.	This presents the same issue as '278 Patent claim 2. *See* discussion above.
'278 Patent claim 17	The method of claim 16, further comprising: generating said output morph weight set stream by interpolating between morph weight sets at said transition start time and said transition end time according to a desired frame rate of said output sequence of animated characters	Such interpolation was used in the prior art. '278 Patent 2:29-32.

5. The Draftsman's Art

This case illustrates the danger that exists when the novel portions of an invention are claimed too broadly. As noted above, the claims here are drafted to give the impression of tangibility, but the Supreme Court has "long warn[ed] . . .

against interpreting § 101 in ways that make patent eligibility depend simply on the draftsman's art. " *Alice*, 134 S. Ct. at 2351 (citing *Mayo*, 132 S.Ct. at 1294). When examined in light of the prior art, the claims are directed to an abstract idea, and lack an "inventive concept" "sufficient to ensure that the patent in practice amounts to significantly more than a patent upon the [abstract idea] itself." *Id*. at 2355 (citations omitted).

IV. **Conclusion**

For the foregoing reasons, the Court would GRANT the Motion, and hold '576 Patent claims 1, 7-9, and 13, and '278 Patent claims 1-4, 6, 9, 13, and 15-17 invalid under 35 U.S.C. § 101.

Dated: This 22nd day of September, 2014.

George H. Wu

GEORGE H. WU
United States District Judge

UNITED STATES DISTRICT COURT
CENTRAL DISTRICT OF CALIFORNIA
WESTERN DIVISION

McRo, Inc., d.b.a. Planet Blue,

Plaintiff,

v.

Bandai Namco Games America, Inc., et al.

Defendants.

Bandai Namco Games America, Inc., et al.,

Counterclaim-Plaintiffs,

v.

McRo, Inc., d.b.a. Planet Blue,

Counterclaim-Defendants.

CASE No. 12-cv-10322-GW (FFMx)

FINAL JUDGMENT

Honorable George H. Wu

CONSOLIDATED WITH:
12-cv-10323-GW (FFMx)
12-cv-10326-GW (FFMx)
12-cv-10327-GW (FFMx)
12-cv-10329-GW (FFMx)
12-cv-10331-GW (FFMx)
12-cv-10333-GW (FFMx)
12-cv-10335-GW (FFMx)
12-cv-10337-GW (FFMx)
12-cv-10338-GW (FFMx)
12-cv-10341-GW (FFMx)
12-cv-10342-GW (FFMx)
13-cv-01870-GW (FFMx)
13-cv-01874-GW (FFMx)
14-cv-00332-GW (FFMx)
14-cv-00336-GW (FFMx)
14-cv-00352-GW (FFMx)
14-cv-00358-GW (FFMx)
14-cv-00383-GW (FFMx)
14-cv-00389-GW (FFMx)
14-cv-00417-GW (FFMx)
14-cv-00439-GW (FFMx)

WHEREAS, Defendants' Motion for Judgment on the Pleadings Based on Unpatentability under 35 U.S.C. § 101 (the "Motion") was fully briefed and oral argument heard by the Court on September 18, 2014;

WHEREAS, the Court issued an order granting Defendants' Motion on September 22, 2014, holding that the asserted claims of U.S. Patent No. 6,307,576 (*i.e.* claims 1, 7, 8, 9, and 13 – collectively, "the Asserted Claims of U.S. Patent No. 6,307,576") and the asserted claims of U.S. Patent No. 6,611,278 (*i.e.* claims 1, 2, 3, 4, 6, 9, 13, 15, 16, and 17 – collectively, "the Asserted Claims of U.S. Patent No. 6,611,278") are invalid under 35 U.S.C. § 101;

WHEREAS, in light of the Court's Order granting Defendants' Motion, final judgment should be entered in favor of Defendants and against Plaintiff and Counterclaim-Defendant McRo, Inc., d.b.a. Planet Blue ("Plaintiff").

It is **ADJUDGED** that:

- The Asserted Claims of U.S. Patent No. 6,307,576 are found to be invalid based on unpatentability under 35 U.S.C. § 101.
- The Asserted Claims of U.S. Patent No. 6,611,278 are found to be invalid based on unpatentability under 35 U.S.C. § 101.

Accordingly, it is **ADJUDGED** that Plaintiff and Counterclaim-Defendant McRo, Inc., d.b.a. Planet Blue ("Plaintiff") takes nothing from Defendants and Counterclaim-Plaintiffs Bandai Namco Games America, Inc.; Sega of America, Inc.; Electronic Arts Inc.; Disney Interactive Studios, Inc.; Capcom USA, Inc.; Neversoft Entertainment, Inc.; Treyarch Corporation; Warner Bros. Interactive Entertainment; LucasArts; Activision Publishing, Inc.; Blizzard Entertainment, Inc.; Infinity Ward, Inc.; Atlus U.S.A., Inc.; Konami Digital Entertainment, Inc.; Square Enix, Inc.; Obsidian Entertainment, Inc.; Naughty Dog, Inc.; Sony Computer Entertainment America LLC; Sucker Punch Productions LLC; Codemasters USA Group, Inc.;

Codemasters, Inc.; The Codemasters Software Company Limited; and Valve Corporation ("Defendants").

All remaining pending motions are **DENIED** as moot.

As Defendants are the prevailing parties in this action, Defendants' costs of court shall be taxed against Plaintiff.

Dated: October 31, 2014

George H. Wu

Hon. George H. Wu,
United States District Judge

US006307576B1

(12) **United States Patent**
Rosenfeld

(10) **Patent No.: US 6,307,576 B1**
(45) **Date of Patent: *Oct. 23, 2001**

(54) **METHOD FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF ANIMATED CHARACTERS**

(76) Inventor: **Maury Rosenfeld**, 1040 N. Las Palmas Ave. No. 25, Los Angeles, CA (US) 90038

(*) Notice: This patent issued on a continued prosecution application filed under 37 CFR 1.53(d), and is subject to the twenty year patent term provisions of 35 U.S.C. 154(a)(2).

Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **08/942,987**

(22) Filed: **Oct. 2, 1997**

(51) **Int. Cl.**[7] **G06T 15/70**

(52) **U.S. Cl.** **345/956**; 345/951; 345/955; 345/473

(58) **Field of Search** 345/473, 951, 345/953, 956, 957, 955

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,600,919	*	7/1986	Stern	345/473
4,884,972	*	12/1989	Gasper et al.	345/473
5,111,409	*	5/1992	Gasper et al.	345/302
5,416,899	*	5/1995	Poggio et al.	345/475
5,613,056	*	3/1997	Gasper et al.	345/473
5,657,426	*	8/1997	Waters et al.	704/276
5,663,517	*	9/1997	Oppenheim	84/649
5,684,942	*	11/1997	Kimura	345/473
5,692,117	*	11/1997	Berrend et al.	345/475
5,717,848	*	2/1998	Watanabe et al.	345/474
5,818,461	*	3/1999	Rouet et al.	345/473
5,880,788	*	3/1999	Bregler	348/515
5,907,351	*	5/1999	Chen et al.	348/14
6,097,381	*	8/2000	Scott et al.	345/302
6,108,011	*	8/2000	Fowler	345/441
6,147,692	*	11/2000	Shaw et al.	345/433
6,232,965	*	5/2001	Scott et al.	707/500

OTHER PUBLICATIONS

Beier et al; Feature–Based Image Metamorphosis; Computer Graphics, 26, 2, Jul. 1992.*

Brooke et al; Computer graphics animations of talking faces based on stochastic models; Proceedings; ISSIPNN 1994 International Symposium; p. 73–76 vol. 1, Apr. 1994.*

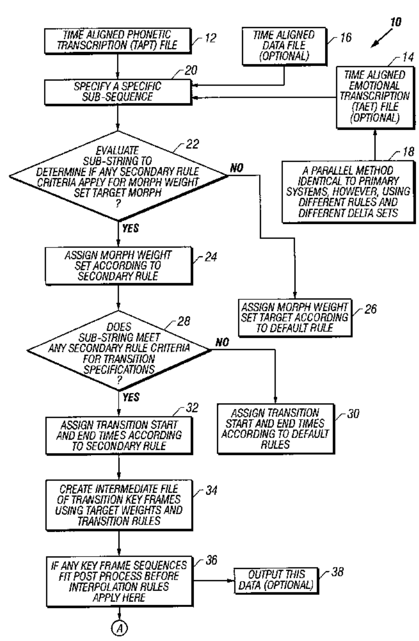
* cited by examiner

Primary Examiner—Matthew Luu
Assistant Examiner—Ryan Yang
(74) *Attorney, Agent, or Firm*—The Hecker Law Group

(57) **ABSTRACT**

A method for controlling and automatically animating lip synchronization and facial expressions of three dimensional animated characters using weighted morph targets and time aligned phonetic transcriptions of recorded text. The method utilizes a set of rules that determine the systems output comprising a stream of morph weight sets when a sequence of timed phonemes and/or other timed data is encountered. Other data, such as timed emotional state data or emotemes such as “surprise, “disgust, “embarrassment”, “timid smile”, or the like, may be inputted to affect the output stream of morph weight sets, or create additional streams.

26 Claims, 4 Drawing Sheets



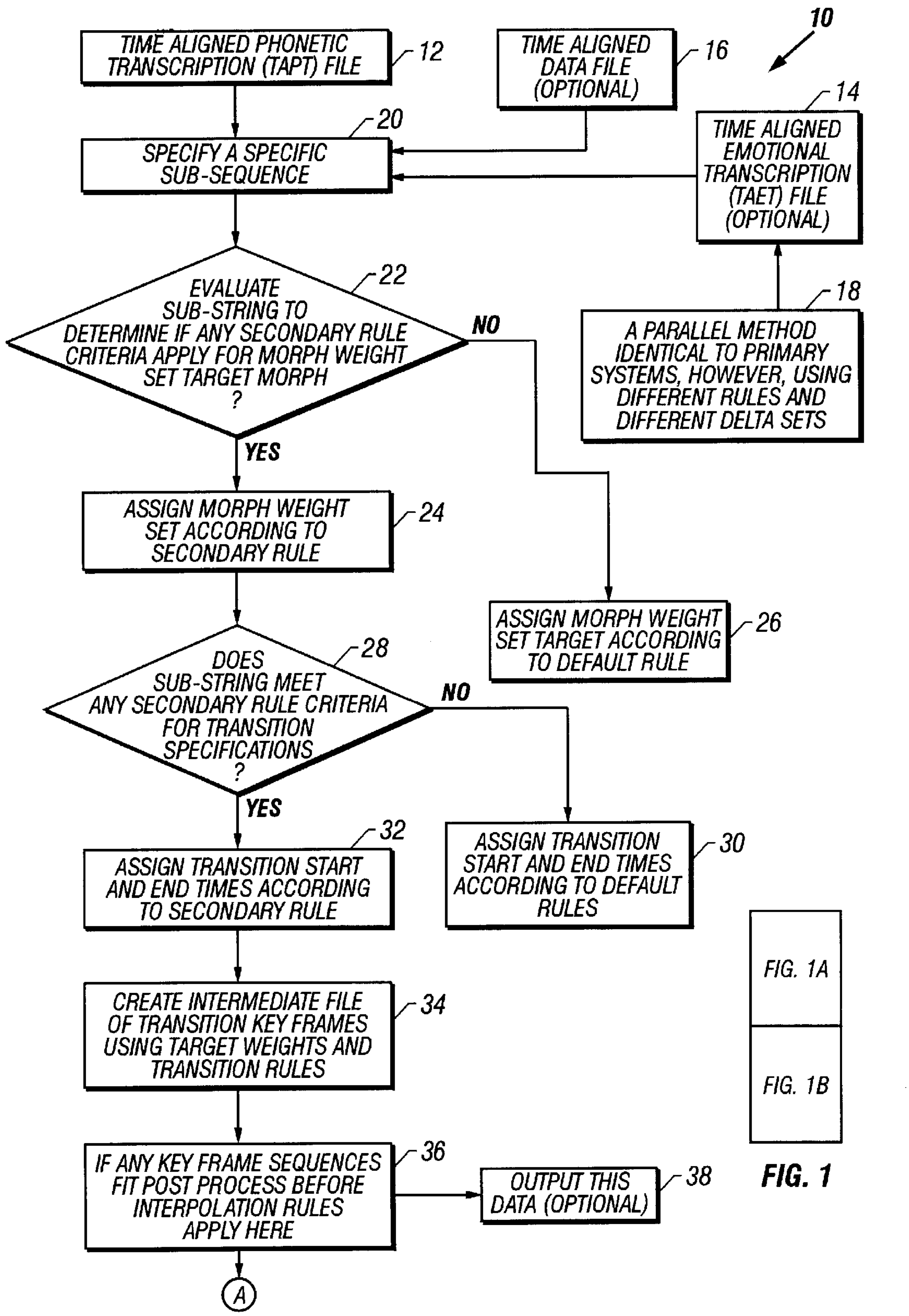

10
TIME ALIGNED PHONETIC TRANSCRIPTION (TAPT) FILE
12
TIME ALIGNED DATA FILE (OPTIONAL)
16
14
TIME ALIGNED EMOTIONAL TRANSCRIPTION (TAET) FILE (OPTIONAL)
20
SPECIFY A SPECIFIC SUB-SEQUENCE
22
EVALUATE SUB-STRING TO DETERMINE IF ANY SECONDARY RULE CRITERIA APPLY FOR MORPH WEIGHT SET TARGET MORPH ?
NO
YES
18
A PARALLEL METHOD IDENTICAL TO PRIMARY SYSTEMS, HOWEVER, USING DIFFERENT RULES AND DIFFERENT DELTA SETS
ASSIGN MORPH WEIGHT SET ACCORDING TO SECONDARY RULE
24
ASSIGN MORPH WEIGHT SET TARGET ACCORDING TO DEFAULT RULE
26
28
DOES SUB-STRING MEET ANY SECONDARY RULE CRITERIA FOR TRANSITION SPECIFICATIONS ?
NO
YES
32
ASSIGN TRANSITION START AND END TIMES ACCORDING TO SECONDARY RULE
ASSIGN TRANSITION START AND END TIMES ACCORDING TO DEFAULT RULES
30
CREATE INTERMEDIATE FILE OF TRANSITION KEY FRAMES USING TARGET WEIGHTS AND TRANSITION RULES
34
IF ANY KEY FRAME SEQUENCES FIT POST PROCESS BEFORE INTERPOLATION RULES APPLY HERE
36
OUTPUT THIS DATA (OPTIONAL)
38
A
FIG. 1A
FIG. 1B
FIG. 1

FIG. 1A

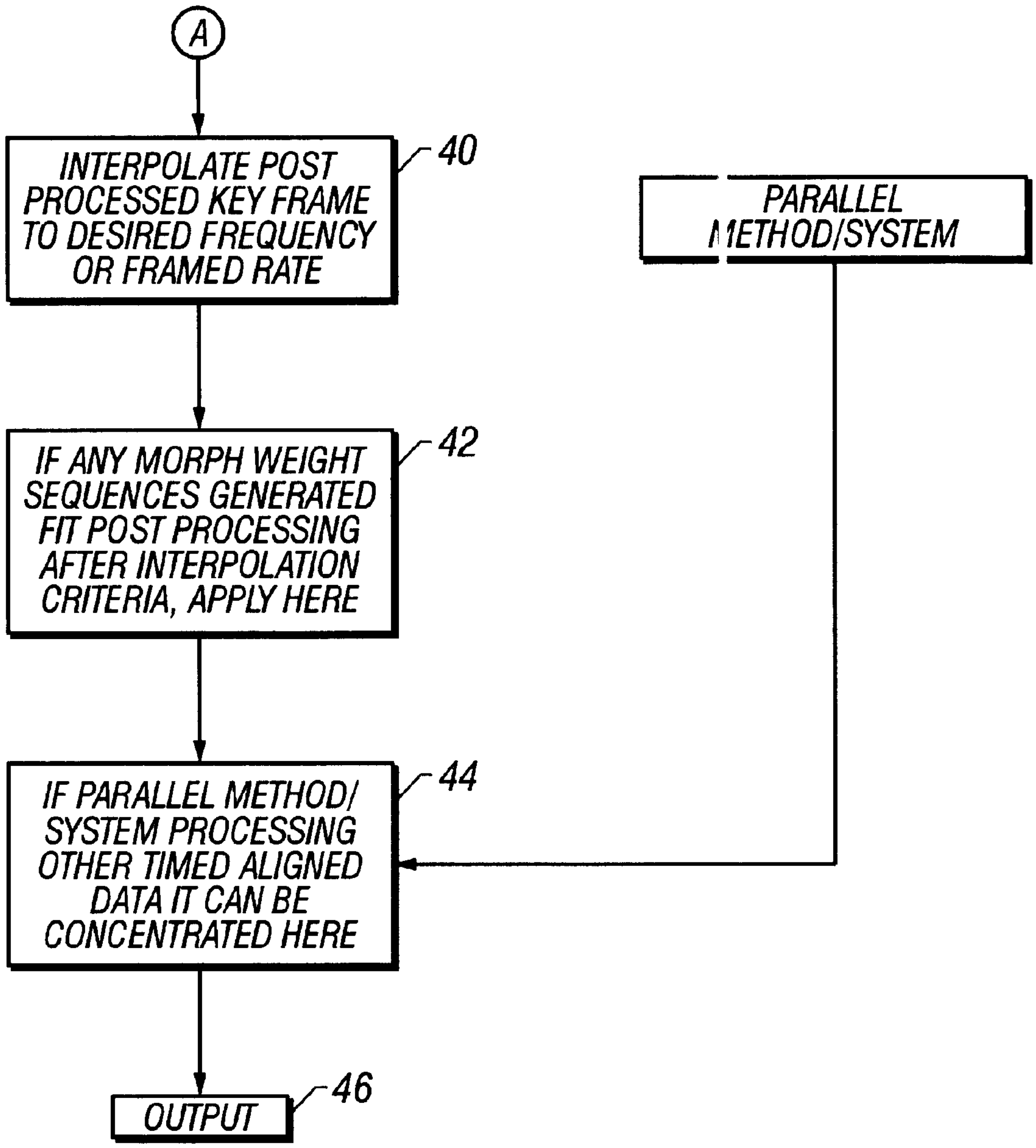

A
INTERPOLATE POST PROCESSED KEY FRAME TO DESIRED FREQUENCY OR FRAMED RATE
40
PARALLEL METHOD/SYSTEM
IF ANY MORPH WEIGHT SEQUENCES GENERATED FIT POST PROCESSING AFTER INTERPOLATION CRITERIA, APPLY HERE
42
IF PARALLEL METHOD/ SYSTEM PROCESSING OTHER TIMED ALIGNED DATA IT CAN BE CONCENTRATED HERE
44
OUTPUT
46

FIG. 1B

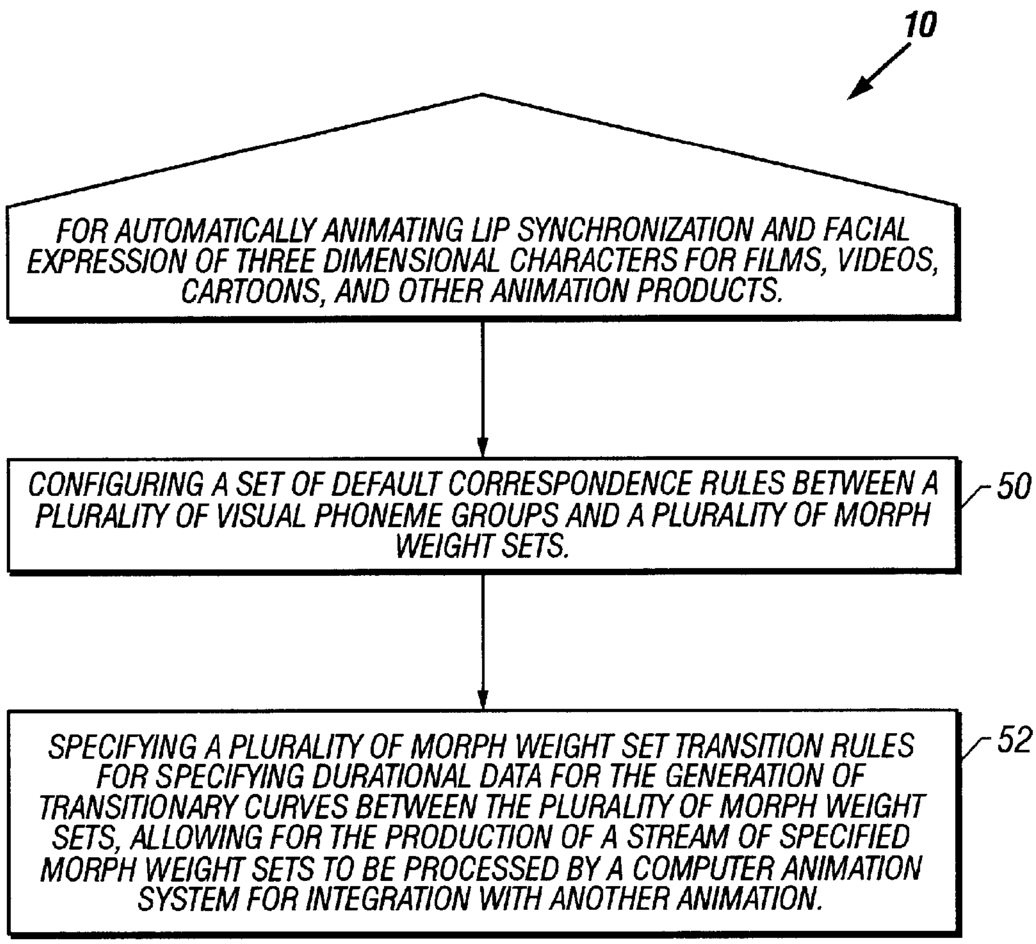

10
FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF THREE DIMENSIONAL CHARACTERS FOR FILMS, VIDEOS, CARTOONS, AND OTHER ANIMATION PRODUCTS.
CONFIGURING A SET OF DEFAULT CORRESPONDENCE RULES BETWEEN A PLURALITY OF VISUAL PHONEME GROUPS AND A PLURALITY OF MORPH WEIGHT SETS.
50
SPECIFYING A PLURALITY OF MORPH WEIGHT SET TRANSITION RULES FOR SPECIFYING DURATIONAL DATA FOR THE GENERATION OF TRANSITIONARY CURVES BETWEEN THE PLURALITY OF MORPH WEIGHT SETS, ALLOWING FOR THE PRODUCTION OF A STREAM OF SPECIFIED MORPH WEIGHT SETS TO BE PROCESSED BY A COMPUTER ANIMATION SYSTEM FOR INTEGRATION WITH ANOTHER ANIMATION.
52

FIG. 2

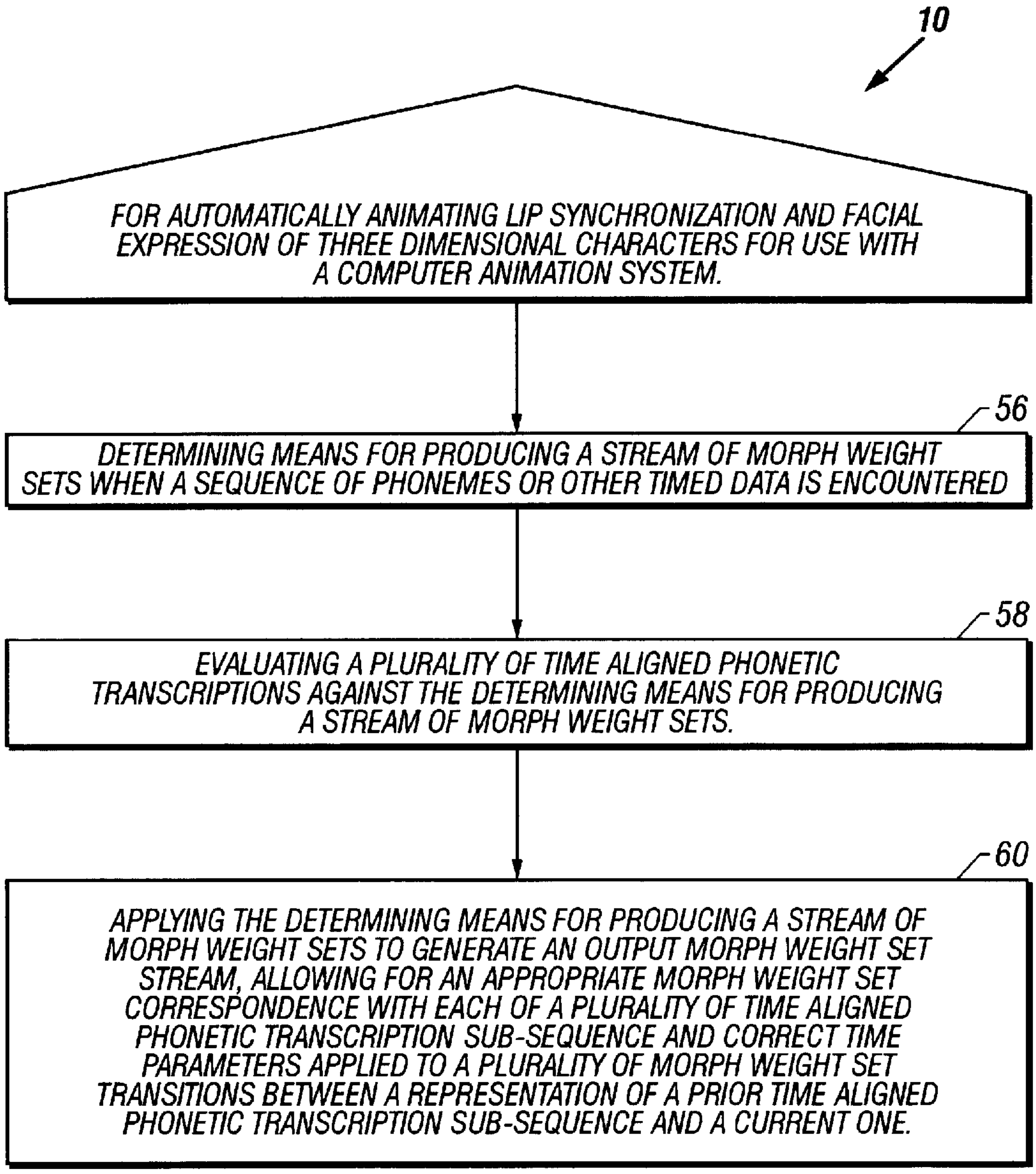

10
FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF THREE DIMENSIONAL CHARACTERS FOR USE WITH A COMPUTER ANIMATION SYSTEM.
56
DETERMINING MEANS FOR PRODUCING A STREAM OF MORPH WEIGHT SETS WHEN A SEQUENCE OF PHONEMES OR OTHER TIMED DATA IS ENCOUNTERED
58
EVALUATING A PLURALITY OF TIME ALIGNED PHONETIC TRANSCRIPTIONS AGAINST THE DETERMINING MEANS FOR PRODUCING A STREAM OF MORPH WEIGHT SETS.
60
APPLYING THE DETERMINING MEANS FOR PRODUCING A STREAM OF MORPH WEIGHT SETS TO GENERATE AN OUTPUT MORPH WEIGHT SET STREAM, ALLOWING FOR AN APPROPRIATE MORPH WEIGHT SET CORRESPONDENCE WITH EACH OF A PLURALITY OF TIME ALIGNED PHONETIC TRANSCRIPTION SUB-SEQUENCE AND CORRECT TIME PARAMETERS APPLIED TO A PLURALITY OF MORPH WEIGHT SET TRANSITIONS BETWEEN A REPRESENTATION OF A PRIOR TIME ALIGNED PHONETIC TRANSCRIPTION SUB-SEQUENCE AND A CURRENT ONE.

FIG. 3

METHOD FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF ANIMATED CHARACTERS

BACKGROUND OF THE INVENTION

1. Field of Invention

This invention relates generally to animation producing methods and apparatuses, and more particularly is directed to a method for automatically animating lip synchronization and facial expression for three dimensional characters.

2. Description of the Related Art

Various methods have been proposed for animating lip synchronization and facial expressions of animated characters in animated products such as movies, videos, cartoons, CD's, and the like. Prior methods in this area have long suffered from the need of providing an economical means of animating lip synchronization and character expression in the production of animated products due to the extremely laborious and lengthy protocols of such prior traditional and computer animation techniques. These shortcomings have significantly limited all prior lip synchronization and facial expression methods and apparatuses used for the production of animated products. Indeed, the limitations of cost, time required to produce an adequate lip synchronization or facial expression in an animated product, and the inherent limitations of prior methods and apparatuses to satisfactorily provide lip synchronization or express character feelings and emotion, leave a significant gap in the potential of animated methods and apparatuses in the current state of the art.

Time aligned phonetic transcriptions (TAPTS) are a phonetic transcription of a recorded text or soundtrack, where the occurrence in time of each phoneme is also recorded. A "phonemes" is defined as the smallest unit of speech, and corresponds to a single sound. There are several standard phonetic "alphabets" such as the International Phonetic Alphabet, and TIMIT created by Texas Instruments, Inc. and MIT. Such transcriptions can be created by hand, as they currently are in the traditional animation industry and are called "x" sheets, or "gray sheets" in the trade. Alternatively such transcriptions can be created by automatic speech recognition programs, or the like.

The current practice for three dimensional computer generated speech animation is by manual techniques commonly using a "morph target" approach. In this practice a reference model of a neutral mouth position, and several other mouth positions, each corresponding to a different phoneme or set of phonemes is used. These models are called "morph targets". Each morph target has the same topology as the neutral model, the same number of vertices, and each vertex on each model logically corresponds to a vertex on each other model. For example, vertex #n on all models represents the left corner of the mouth, and although this is the typical case, such rigid correspondence may not be necessary.

The deltas of each vertex on each morph target relative to the neutral are computed as a vector from each vertex n on the reference to each vertex n on each morph target. These are called the delta sets. There is one delta set for each morph target.

In producing animation products, a value usually from 0 to 1 is assigned to each delta set by the animator and the value is called the "morph weight". From these morph weights, the neutral's geometry is modified as follows: Each vertex N on the neutral has the corresponding delta set's vertex multiplied by the scalar morph weight added to it. This is repeated for each morph target, and the result summed. For each vertex v in the neutral model:

$$|\text{result}| = |\text{neutral}| + \sum_{x=1}^{n} |\text{delta set}_x| * \text{morph weight}_x$$

|delta setx|*morph weightx

where the symbol |xxx| is used to indicate the corresponding vector in each referenced set. For example, Iresult is the corresponding resultant vertex to vertex v in the neutral model |neutral| and |delta setx| is the corresponding vector for delta set x.

If the morph weight of the delta set corresponding to the morph target of the character saying, for example, the "oh" sound is set to 1, and all others are set to 0, the neutral would be modified to look like the "oh target. If the situation was the same, except that the "oh" morph weight was 0.5, the neutral's geometry is modified half way between neutral and the "oh" morph target.

Similarly, if the situation was as described above, except "oh" weight was 0.3 and the "ee" morph weight was at 0.7, the neutral geometry is modified to have some of the "oh" model characteristics and more of the "ee" model characteristics. There also are prior blending methods including averaging the delta sets according to their weights.

Accordingly, to animate speech, the artist needs to set all of these weights at each frame to an appropriate value. Usually this is assisted by using a "keyframe" approach, where the artist sets the appropriate weights at certain important times ("keyframes") and a program interpolates each of the channels at each frame. Such keyframe approach is very tedious and time consuming, as well as inaccurate due to the large number of keyframes necessary to depict speech.

The present invention overcomes many of the deficiencies of the prior art and obtains its objectives by providing an integrated method embodied in computer software for use with a computer for the rapid, efficient lip synchronization and manipulation of character facial expressions, thereby allowing for rapid, creative, and expressive animation products to be produced in a very cost effective manner.

Accordingly, it is the primary object of this invention to provide a method for automatically animating lip synchronization and facial expression of three dimensional characters, which is integrated with computer means for producing accurate and realistic lip synchronization and facial expressions in animated characters. The method of the present invention further provides an extremely rapid and cost effective means to automatically create lip synchronization and facial expression in three dimensional animated characters.

Additional objects and advantages of the invention will be set forth in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The objects and advantages of the invention may be realized and obtained by means of the instrumentalities and combinations particularly pointed out in the appended claims.

SUMMARY OF THE INVENTION

To achieve the foregoing objects, and in accordance with the purpose of the invention as embodied and broadly described herein, a method is provided for controlling and automatically animating lip synchronization and facial

expressions of three dimensional animated characters using weighted morph targets and time aligned phonetic transcriptions of recorded text, and other time aligned data. The method utilizes a set of rules that determine the systems output comprising a stream or streams of morph weight sets when a sequence of timed phonemes or other timed data is encountered. Other timed data, such as pitch, amplitued, noise amounts, or emotional state data or emotemes such as "surprise, "disgust, "embarrassment", "timid smile", or the like, may be inputted to affect the output stream of morph weight sets.

The methodology herein described allows for automatically animating lip synchronization and facial expression of three dimensional characters in the creation of a wide variety of animation products, including but not limited to movies, videos, cartoons, CD's, software, and the like. The method and apparatuses herein described are operably integrated with computer software and hardware.

In accordance with the present invention there also is provided a method for automatically animating lip synchronization and facial expression of three dimensional characters for films, videos, cartoons, and other animation products, comprising configuring a set of default correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and specifying a plurality of morph weight set transition rules for specifying durational data for the generation of transitionary curves between the plurality of morph weight sets, allowing for the production of a stream of specified morph weight sets to be processed by a computer animation system for integration with other animation, whereby animated lip synchronization and facial expression of animated characters may be automatically controlled and produced.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate a preferred embodiment of the invention and, together with a general description given above and the detailed description of the preferred embodiment given below, serve to explain the principles of the invention.

FIG. **1** is a flow chart showing the method of the invention with an optional time aligned emotional transcription file, and another parallel timed data file, according to the invention.

FIG. **2** is a flow chart illustrating the principal steps of the present method, according to the invention.

FIG. **3** is another representational flow chart illustrating the present method, according to the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Reference will now be made in detail to the present preferred embodiments of the invention as illustrated in the accompanying drawings.

In accordance with the present invention, there is provided as illustrated in FIGS. **1**–**3**, a method for controlling and automatically animating lip synchronization and facial expressions of three dimensional animated characters using weighted morph targets and time aligned phonetic transcriptions of recorded text. The method utilizes a set of rules that determine the systems output comprising a stream of morph weight sets when a sequence of timed phonemes is encountered. Other timed data, such as timed emotional state data or emotemes such as "surprise, "disgust, "embarrassment", "timid smilen", pitch, amplitued, noise amounts or the like, may be inputted to affect the output stream of morph weight sets.

The method comprises, in one embodiment, configuring a set of default correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and specifying a plurality of morph weight set transition rules for specifying durational data for the generation of transitionary curves between the plurality of morph weight sets, allowing for the production of a stream of specified morph weight sets to be processed by a computer animation system for integration with other animation, whereby animated lip synchronization and facial expression of animated characters may be automatically produced.

There is also provided, according to the invention a method for automatically animating lip synchronization and facial expression of three dimensional characters for use with a computer animation system, comprising the steps of: determining means for producing a stream of morph weight sets when a sequence of phonemes is encountered; evaluating a plurality of time aligned phonetic transcriptions or other timed data such as pitch, amplitude, noise amounts and the like, against the determining means for producing a stream of morph weight sets; applying said determining means for producing a stream of morph weight sets to generate an output morph weight set stream, allowing for an appropriate morph weight set correspondence with each of a plurality of time aligned phonetic transcription sub-sequences and correct time parameters applied to a plurality of morph weight set transitions between a representation of a prior time aligned phonetic transcription sub-sequence and a current one, whereby lip synchronization and facial expressions of animated characters is automatically controlled and produced.

The method preferably comprises a set of rules that determine what the output morph weight set steam will be when any sequence of phonemes and their associated times is encountered. As used herein, a "morph weight set" is a set of values, one for each delta set, that, when applied as described, transform the neutral model to some desired state, such as speaking the "oo" sound or the "th" sound. Preferably, one model id designated as the anchor model, which the deltas are computed in reference to. If for example, the is a morph target that represents all possible occurrences of an "e" sound perfectly, it's morph weight set would be all zeros for all delta sets except for the delta set corresponding to the "ee" morph target, which would set to 1.

Preferably, each rule comprises two parts, the rule's criteria and the rule's function. Each sub-sequence of time aligned phonetic transcription (TAPT) or other timed data such as pitch, amplitude, noise amount or the like, is checked against a rule's criteria to see if that rule is applicable. If so, the rule's function is applied to generate the output. The primary function of the rules is to determined 1) the appropriate morph weight set correspondence with each TAPT sub-sequence; and 2) the time parameters of the morph weight set transitions between the representation of the prior TAPT sub-sequence or other timed data,and the current one. Conditions 1) and 2) must be completely specified for any sequence of phonemes and times encountered. Together, such rules are used to create a continuous stream of morph weight sets.

In the present method, it is allowable for more than one phoneme to be represented by the same morph target, for example, "sss" and "zzz". Visually, these phonemes appear

similar. Through the use of such rules, the user can group phonemes together that have a similar visual appearance into visual phonemes"that function the same as one another. It is also acceptable, through the rules, to ignore certain phoneme sequences. For example, a rule could specify: "If in the TAPT, there are two or more adjacent phonemes that are in the same "visual phoneme" group, all but the first are ignored".

The rules of the present method may be categorized in three main groupings; default rules, auxiliary rules and post processing rules. The default rules must be complete enough to create valid output for any TAPT encountered at any point in the TAPT. The secondary rules are used in special cases; for example, to substitute alternative morph weight set correspondences and/or transition rules if the identified criteria are met. The post processing rules are used to further manipulate the morph weight set stream after the default or secondary rules are applied, and can further modify the members of the morph weight sets determined by the default and secondary rules and interpolation.

If for example, a specific TAPT sub-sequence does not fit the criteria for any secondary rules, then the default rules take effect. If, on the other hand, the TAPT sub-sequence does fit the criteria for a secondary rule(s) they take precedence over the default rules. A TAPT sub-sequence take into account the current phoneme and duration, and a number of the preceding and following phonemes and duration's as well may be specified.

Preferably, the secondary rules effect morph target correspondence and weights, or transition times, or both. Secondary rules can create transitions and correspondences even where no phoneme transitions exist. The secondary rules can use as their criteria the phoneme, the duration or the phoneme's context in the output stream, that is what phonemes are adjacent or in the neighborhood to the current phoneme, what the adjacent durations are, and the like.

The post processing rules are preferably applied after a preliminary output morph weight set is calculated so as to modify it. Post processing rules can be applied before interpolation and/or after interpolation, as described later in this document. Both the secondary and post processing rules are optional, however, they may in certain applications be very complex, and in particular circumstances contribute more to the output than the default rules.

In FIG. **1**, a flow chart illustrates the preferred steps of the methodology **10** for automatically animating lip synchronization and facial expression of three dimensional animated characters of the present invention. A specific sub-sequence **20** is selected form the TAPT file **12** and is evaluated **22** to determine if any secondary rule criteria for morph weight set target apply. Time aligned emotional transcription file **14** data may be inputted or data from an optional time aligned data file **16** may be used. Also shown is a parallel method **18** which may be configured identical to the primary method described, however, using different timed data rules and different delta sets. Sub-sequence **20** is evaluated **22** to determine if any secondary rule criteria apply. If yes, then a morph weight set is assigned **24** according to the secondary rules, if no, then a morph weight set is assigned **26** according to the default rules. If the sub-string meets any secondary rule criteria for transition specification **28** then a transition start and end time are assigned according to the secondary rules **32**, if no, then assign transition start and end times **30** according to default rules. Then an intermediate file of transition keyframes using target weights and transition rules as generated are created **34**, and if any keyframe sequences fit post process before interpolation rules they are applied here **36**. This data may be output **38** here if desired. If not, then interpolate using any method post processed keyframes to a desired frequency or frame rate **40** and if any morph weight sequences generated fit post processing after interpolation criteria, they are applied **42** at this point. If parallel methods or systems are used to process other timed aligned data, they may be concatenated here **44**, and the data output **46**.

In FIG. **2**, the method for automatically animating lip synchronization and facial expression of three dimensional characters for films, videos, cartoons, and other animation products **10** is shown according to the invention, where box **50** show the step of configuring a set of default correspondence rules between a plurality of visual phoneme groups or other timed input data and a plurality of morph weight sets. Box **52** shows the steps of specifying a plurality of morph weight set transition rules for specifying durational data for the generation of transitionary curves between the plurality of morph weight sets, allowing for the production of a stream of specified morph weight sets to be processed by a computer animation system for integration with other animation, whereby animated lip synchronization and facial expression of animated characters may be automatically produced.

With reference now to FIG. **3**, method **10** for automatically animating lip synchronization and facial expression of three dimensional characters for use with a computer animation system is shown including box **56** showing the step of determining means for producing a stream of morph weight sets when a sequence of phonemes is encountered. Box **58**, showing the step of evaluating a plurality of time aligned phonetic transcriptions or other timed at such as pitch, amplitude, noise amounts, and the like, against said determining means for producing a stream of morph weight sets. In box **60** the steps of applying said determining means for producing a stream of morph weight sets to generate an output morph weight set stream, allowing for an appropriate morph weight set correspondence with each of a plurality of time aligned phonetic transcription sub-sequences and correct time parameters applied to a plurality of morph weight set transitions between a representation of a prior time aligned phonetic transcription sub-sequence and a current one, whereby lip synchronization and facial expressions of animated characters is automatically controlled and produced are shown according to the invention.

In operation and use, the user must manually set up default correspondence rules between all visual phoneme groups and morph weight sets. To do this, the user preferably specifies the morph weight sets which correspond to the model speaking, for example the "oo" sound, the "th" sound, and the like. Next, default rules must be specified. These rules specify the durational information needed to generate appropriate transitionary curves between morph weight sets, such as transition start and end times. A "transition" between two morph weigh sets is defined as each member of the morph weight set transitions from it's current state to it's target state, starting at the transition start time and ending at the transition end time. The target state is the morph weight set determined by a correspondence rule.

The default correspondence rules and the default morph weight set transition rules define the default system behavior. If all possible visual phoneme groups or all members of alternative data domains have morph weight set correspondence, any phoneme sequence can be handled with this rule set alone. However, additional rules are desirable for effects, exceptions, and uniqueness of character, as further described below.

According to the method of the invention, other rules involving phoneme's duration and/or context can be specified. Also, any other rules that do not fit easily into the above mentioned categories can be specified. Examples of such rules are described in greater detail below and are termed the "secondary rules". If a timed phoneme or sub-sequence of timed phonemes do not fit the criteria for any of the secondary rules, the default rules are applied as seen in FIG. **1**.

It is seen that through the use of these rules, an appropriate morph weight stream is produced. The uninterpolated morph weight stream has entries only at transition start and end time, however. These act as keyframes. A morph weight set may be evaluated at any time by interpolating between these keyframes, using conventional methods. This is how the output stream is calculated each desired time frame. For example, for television productions, the necessary resolution is 30 evaluations per second.

The post processing rules may be applied either before or after the above described interpolation step, or both. Some rules may apply only to keyframes before interpolation, some to interpolated data. If applied before the interpolation step, this affects the keyframes. if applied after, it effects the interpolated data. Post processing can use the morph weight sets calculated by the default and secondary rules. Post processing rules can use the morph weigh sets or sequences as in box **44** of FIG. **1**, calculated by the default and secondary rules. Post processing rules can modify the individual members of the morph weight sets previously generated. Post processing rules may be applied in addition to other rules, including other post processing rules. Once the rule set up is completed as described, the method of the present invention can take any number and length TAPT's as input, and automatically output the corresponding morph weight set stream as seen in FIGS. **1–3**.

For example, a modeled neutral geometric representation of a character for an animated production such as a movies, video, cartoon, CD or the like, with six morph targets, and their delta sets determined. Their representations, for example, are as follows:

Delta Set	Visual Representation
1	"h"
2	"eh"
3	"1"
4	"oh"
5	exaggerated "oh"
6	special case "eh" used during a "snide laugh" sequences

In this example, the neutral model is used to represent silence. The following is an example of a set of rules, according to the present method, of course this is only an example of a set of rules which could be use for illustrative purposes, and many other rules could be specified according to the method of the invention.

Default Rules

Default Correspondence Rules;

Criteria: Encounter a "h" as in "house"

Function: Use morph weight set (1,0,0,0,0,0) as transition target.

Criteria: Encounter an "eh" as in "bet"

Function: Use morph weight set (0,1,0,0,0,0) as transition target.

Criteria: Encounter a "1" as in "old"

Function: Use morph weight set (0,0,1,0,0,0) as transition target.

Criteria: Encounter an "oh" as in "old"

Function: Use morph weight set (0,0,0,1,0,0) as transition target.

Criteria: encounter a "silence"

Function: use morph weight set (0,0,0,0,0,0) as transition target.

Default Transition Rule:

Criteria: Encounter any phoneme

Function: Transition start time=(the outgoing phoneme's end time)−0.1*(the outgoing phoneme's duration);

transition end time=(the incoming phoneme's start time)+0.1* (the incoming phoneme's duration)

Secondary Rules

Criteria: Encounter an "oh" with a duration greater than 1.2 seconds.

Function: Use morph weigh set (0,0,0,0,1,0)

Criteria: Encounter an "eh" followed by an "oh" and preceded by an "h".

Function: Use morph weigh set (0,0,0,0,0,1) as transition target.

Criteria: Encounter any phoneme preceded by silence

Function: Transition start time=(the silence's end time)−0.1*(the incoming phoneme's duration):Transition end time=the incoming phoneme's start time

Criteria: Encounter silence preceded by any phoneme.

Function: Transition start time=the silence's start time +0.1* (the outgoing phoneme's duration)

Post Processing Rules

Criteria: Encounter a phoneme duration under 0.22 seconds.

Function: Scale the transition target determined by the default and secondary rules by 0.8 before interpolation.

Accordingly, using this example, if the user were to use these rules for the spoken word "Hello", at least four morph targets and a neutral target would be required, that is, one each for the sound of "h", "e","1","oh" and their associated delta sets. For example, a TAPT representing the spoken word "hello" could be configured as,

Time	Phoneme
0.0	silence begins
0.8	silence ends, "h" begins
1.0	"h" ends, "eh" begins
1.37	"eh" ends, "1" begins
1.6	"1" ends, "oh" begins
2.1	"oh" ends, silence begins.

The method, for example embodied in computer software for operation with a computer or computer animation system would create an output morph weight set stream as follows:

Time	D.S.1 ("h")	D.S.2 ("eh")	D.S.3 ("1")	D.S.4 ("oh")	D.S.5 (aux"oh")	D.S.6
0.0	0	0	0	0	0	0
0.78	0	0	0	0	0	0
0.8	1	0	0	0	0	0
0.98	1	0	0	0	0	0
1.037	0	1	0	0	0	0
1.333	0	1	0	0	0	0
1.403	0	0	1	0	0	0
1.667	0	0	1	0	0	0

-continued

Time	D.S.1 ("h")	D.S.2 ("eh")	D.S.3 ("1")	D.S.4 ("oh")	D.S.5 (aux"oh")	D.S.6
1.74	0	0	0	1	0	0
2.1	0	0	0	1	0	0
2.14	0	0	0	0	0	0

Such morph weight sets act as keyframes, marking the transitionary points. A morph weight set can be obtained for any time within the duration of the TAPT by interpolating between the morph weight sets using conventional methods well known in the art. Accordingly, a morph weight set can be evaluated at every frame. However, the post processing rules can be applied to the keyframes before interpolation as in box **36** of FIG. **1**, or to the interpolated data as in box **40** of FIG. **1**. From such stream of morph weight sets, the neutral model is deformed as described above, and then sent to a conventional computer animation system for integration with other animation. Alternatively, the morph weight set stream can be used directly by an animation program or package, wither interpolated or not.

The rules of the present invention are extensible and freeform in the sense that they may be created as desired and adapted to a wide variety of animation characters, situations, and products. As each rule comprise a criteria and function, as in an "if . . . then . . . else" construct. The following are illustrative examples of other rules which may be used with the present methodology.

For example, use {0,0,0,0 . . . 0} as the morph weight set when a "m" is encountered. This is a type of default rule, where:

Criteria: Encounter a "m" phoneme of any duration.

Function: Use a morph weight set {0,0,0,0 . . . 0} as a transition target.

Another example would be creating several slightly different morph targets for each phoneme group, and using them randomly each time that phoneme is spoken. This would give a more random, or possibly comical or interesting look to the animation's. This is a secondary rule.

An example of post processing rule, before interpolation would be to add a small amount of random noise to all morph weight channels are all keyframes. This would slightly alter the look of each phoneme to create a more natural look.

Criteria: Encounter any keyframe

Function: Add a small random value to each member of the morph weight set prior to interpolation.

An example of a post processing rule, after interpolation would be to add a component of an auxiliary morph target (one which does not correspond directly to a phoneme) to the output stream in a cyclical manner over time, after interpolation. If the auxiliary morph target had the character's mouth moved to the left, for example, the output animation would have the character's mouth cycling between center to left as he spoke.

Criteria: Encounter any morph weight set generated by interpolation

Function: Add a value calculated through a mathematical expression to the morph weigh set's member that corresponds to the auxiliary morph target's delta set weight. The expression might be, for example: 0.2*sin (0.2*time*2*pi)+0.2. This rule would result in an oscillation of the animated character's mouth every five seconds.

Another example of a secondary rule is to use alternative weight sets(or morph weight set sequences) for certain contexts of phonemes, for example, if an "oh" is both preceded and followed by an "ee" then use an alternate "oh". This type of rule can make speech idiosyncrasies, as well as special sequences for specific words (which are a combination of certain phonemes in a certain context). This type of rule can take into consideration the differences in mouth positions for similar phonemes based on context. For example, the "1" in "hello" is shaped more widely than the "1" in "burly" due to it's proximity to an "eh" as opposed tp a "r".

Criteria: Encounter an "1" preceded by an "r".

Function: Use a specified morph weight set as transition target.

Another secondary rule could be, by way of illustration, that if a phoneme is longer than a certain duration, substitute a different morph target. this can add expressiveness to extended vowel sounds, for instance, if a character says "HELLOOOOOOO!" a more exaggerated "oh" model would be used.

Criteria: Encounter an "oh" longer than 0.5 seconds and less than 1 second.

Function: Use a specified morph weight set as a transition target.

If a phoneme is longer than another phoneme of even longer duration, a secondary rule may be applied to create new transitions between alternate morph targets at certain intervals, which may be randomized, during the phoneme's duration. This will add some animation to extremely long held sounds, avoiding a rigid look. This is another example of a secondary rule

Criteria: Encounter an "oh" longer than 1 second long.

Function: Insert transitions between a defined group of morph weight sets at 0.5 second intervals, with transition duration's of 0.2 seconds until the next "normal" transition start time is encountered.

If a phoneme is shorter than a certain duration, its corresponding morph weight may be scaled by a factor smaller than 1. This would create very short phonemes not appear over articulated. Such a post processing rule, applied before interpolation would comprise:

Criteria: Encounter a phoneme duration shorter than 0.1 seconds.

Function: Multiply all members of the transition target (already determined by default and secondary rules by duration/0.1.

As is readily apparent a wide variety of other rules can be created to add individuality to the different characters.

A further extension of the present method is to make a parallel method or system, as depicted in box **14** of FIG. **1**, that uses time aligned emotional transcriptions (TAET) that correspond to facial models of those emotions. Using the same techniques as previously described additional morph weight set streams can be created that control other aspects of the character that reflect facial display of emotional state. Such morph weight set streams can be concatenated with the lip synchronization stream. In addition, the TAET data can be used in conjunction with the lip synchronization secondary rules to alter the lip synchronization output stream. For example:

Criteria: An "L" is encountered in the TAPT and the nearest "emoteme" in the TAET is a "smile".

Function: Use a specified morph weight set as transition target.

As is evident from the above description, the automatic animation lip synchronization and facial expression method described may be used on a wide variety of animation products. The method described herein provides an

extremely rapid, efficient, and cost effective means to provide automatic lip synchronization and facial expression in three dimensional animated characters. The method described herein provides, for the first time, a rapid, effective, expressive, and inexpensive means to automatically create animated lip synchronization and facial expression in animated characters. The method described herein can create the necessary morph weight set streams to create speech animation when given a time aligned phonetic transcription of spoken text and a set of user defined rules for determining appropriate morph weight sets for a given TAPT sequence. This method also defines rules describing a method of transitioning between these sets through time. The present method is extensible by adding new rules, and other timed data may be supplied, such as time "emotemes" that will effect the output data according to additional rules that take this data into account. In this manner, several parallel systems may be used on different types of timed data and the results concatenated, or used independently. Accordingly, additional advantages and modification will readily occur to those skilled in the art. The invention in its broader aspects is, therefore, not limited to the specific methodological details, representative apparatus and illustrative examples shown and described. Accordingly, departures from such details may be made without departing from the spirit or scope of the applicant's inventive concept.

What is claimed is:

1. A method for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:

obtaining a first set of rules that define output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence;

obtaining a timed data file of phonemes having a plurality of sub-sequences;

generating an intermediate stream of output morph weight sets and a plurality of transition parameters between two adjacent morph weight sets by evaluating said plurality of sub-sequences against said first set of rules;

generating a final stream of output morph weight sets at a desired frame rate from said intermediate stream of output morph weight sets and said plurality of transition parameters; and

applying said final stream of output morph weight sets to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters.

2. The method of claim **1** wherein each of said first set of rules comprises a rule's criteria and a rule's function.

3. The method of claim **2** wherein said evaluating comprises:

checking each sub-sequence of said plurality of sub-sequences for compliance with said rule's criteria; and

applying said rule's function upon said compliance.

4. The method of claim **1** wherein said first set of rules comprises a default set of rules and an optional secondary set of rules, said secondary set of rules having priority over said default set of rules.

5. The method of claim **4** wherein said default set of rules is adequate to create said intermediate stream of output morph weight sets and said plurality of transition parameters between two adjacent morph weight sets for all sub-sequences of phonemes in said timed data file.

6. The method of claim **4** wherein said secondary set of rules are used in special cases to substitute alternate output morph weight sets and/or transition parameters between two adjacent morph weight sets.

7. The method of claim **1** wherein said timed data is a timed aligned phonetic transcriptions data.

8. The method of claim **7** wherein said timed data further comprises time aligned data.

9. The method of claim **7** wherein said timed data further comprises time aligned emotional transcription data.

10. The method of claim **1** wherein each of said plurality of transition parameters comprises a transition start time and a transition end time; and said intermediate stream of output morph weight sets having entries at said transition start time and said transition end time.

11. The method of claim **10** wherein said generating a final stream of output morph weight sets comprises:

obtaining the output morph weight set at a desired time by interpolating between said intermediate stream of morph weight sets at said transition start time and said transition end time, said desired time representing a frame of said final stream of output.

12. The method of claim **11**, further comprising:

applying a second set of rules to said output morph weight set for post processing.

13. The method of claim **1** wherein said first set of rules comprises:

correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and

morph weight set transition rules specifying durational data for generating transitionary curves between morph weight sets.

14. An apparatus for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:

a computer system;

a first set of rules in said computer system, said first set of rules defining output morph weight set stream as a function of phoneme sequence and time of said phoneme sequence;

a timed data file readable by said computer system, said timed data file having phonemes with a plurality of sub-sequences;

means, in said computer system, for generating an intermediate stream of output morph weight sets and a plurality of transition parameters between two adjacent morph weight sets by evaluating said plurality of sub-sequences against said first set of rules;

means, in said computer system, for generating a final stream of output morph weight sets at a desired frame rate from said intermediate stream of output morph weight sets and said plurality of transition parameters; and

means, in said computer system, for applying said final stream of output morph weight sets to a sequence of animated characters to produce lip synchronization and facial expression control of said animated characters.

15. The apparatus of claim **14** wherein each of said first set of rules comprises a rule's criteria and a rule's function.

16. The apparatus of claim **15** wherein said evaluating comprises:

checking each sub-sequence of said plurality of sub-sequences for compliance with said rule's criteria; and

applying said rule's function upon said compliance.

17. The apparatus of claim **14** wherein said first set of rules comprises a default set of rules and an optional secondary set of rules, said secondary set of rules having priority over said default set of rules.

18. The apparatus of claim **17** wherein said default set of rules is adequate to create said intermediate stream of output

morph weight sets and said plurality of transition parameters between two adjacent morph weight sets for all subsequences of phonemes in said timed data file.

19. The apparatus of claim **17** wherein said secondary set of rules are used in special cases to substitute alternate output morph weight sets and/or transition parameters between two adjacent morph weight sets.

20. The apparatus of claim **14** wherein said timed data is a timed aligned phonetic transcriptions data.

21. The apparatus of claim **20** wherein said timed data further comprises time aligned data.

22. The apparatus of claim **20** wherein said timed data further comprises time aligned emotional transcription data.

23. The apparatus of claim **14** wherein each of said plurality of transition parameters comprises a transition start time and a transition end time; and said intermediate stream of output morph weight sets having entries at said transition start time and said transition end time.

24. The apparatus of claim **23** wherein said generating a final stream of output morph weight sets comprises:

obtaining the output morph weight set at a desired time by interpolating between said intermediate stream of morph weight sets at said transition start time and said transition end time, said desired time representing a frame of said final stream of output.

25. The apparatus of claim **24**, further comprising:

means for applying a second set of rules to said output morph weight set for post processing.

26. The apparatus of claim **14** wherein said first set of rules comprises:

correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and

morph weight set transition rules specifying durational data for generating transitionary curves between morph weight sets.

* * * * *

US006611278B2

(12) United States Patent
Rosenfeld

(10) Patent No.: US 6,611,278 B2
(45) Date of Patent: *Aug. 26, 2003

(54) **METHOD FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF ANIMATED CHARACTERS**

(76) Inventor: **Maury Rosenfeld**, 4941 Ambrose Ave, Los Angeles, CA (US) 90027

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **09/960,831**

(22) Filed: **Sep. 21, 2001**

(65) **Prior Publication Data**

US 2002/0101422 A1 Aug. 1, 2002

Related U.S. Application Data

(63) Continuation of application No. 08/942,987, filed on Oct. 2, 1997, now Pat. No. 6,307,576.

(51) **Int. Cl.**[7] **G06T 13/00**; G09G 5/00
(52) **U.S. Cl.** **345/956**; 345/473; 345/646
(58) **Field of Search** 345/473, 646, 345/647, 951, 953, 955, 956, 957

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,600,919 A * 7/1986 Stern 345/473
5,416,899 A * 5/1995 Poggio et al. 345/475
5,630,017 A * 5/1997 Gasper et al. 704/276
5,684,942 A * 11/1997 Kimura 345/473
5,717,848 A * 2/1998 Watanabe et al. 345/474
5,741,136 A * 4/1998 Kirksey et al. 434/169
5,818,461 A * 10/1998 Rouet et al. 345/473
5,880,788 A * 3/1999 Bregler 348/515
5,907,351 A * 5/1999 Chen et al. 348/14.12
6,097,381 A * 8/2000 Scott et al. 707/500.1
6,108,011 A * 8/2000 Fowler 345/441
6,147,692 A * 11/2000 Shaw et al. 345/433
6,232,965 B1 * 5/2001 Scott et al. 707/505
6,307,576 B1 * 10/2001 Rosenfeld 345/700

* cited by examiner

Primary Examiner—Jeffery Brier
Assistant Examiner—Ryan Yang
(74) *Attorney, Agent, or Firm*—The Hecker Law Group

(57) **ABSTRACT**

A method for controlling and automatically animating lip synchronization and facial expressions of three dimensional animated characters using weighted morph targets and time aligned phonetic transcriptions of recorded text. The method utilizes a set of rules that determine the systems output comprising a stream of morph weight sets when a sequence of timed phonemes and/or other timed data is encountered. Other data, such as timed emotional state data or emotemes such as "surprise, "disgust, "embarrassment", "timid smile", or the like, may be inputted to affect the output stream of morph weight sets, or create additional streams.

36 Claims, 4 Drawing Sheets

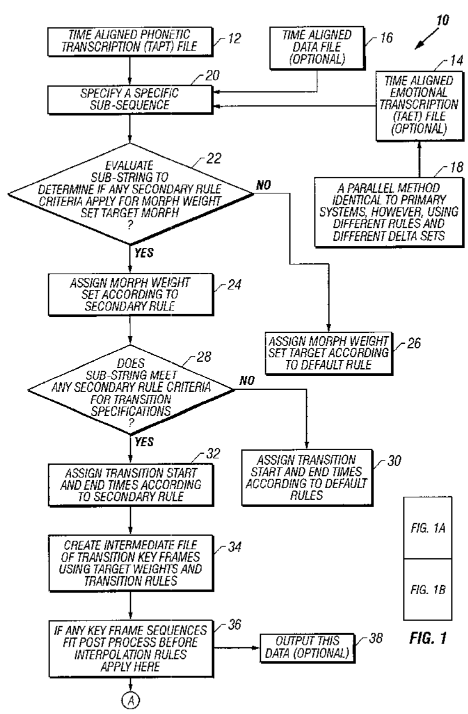


FIG. 1

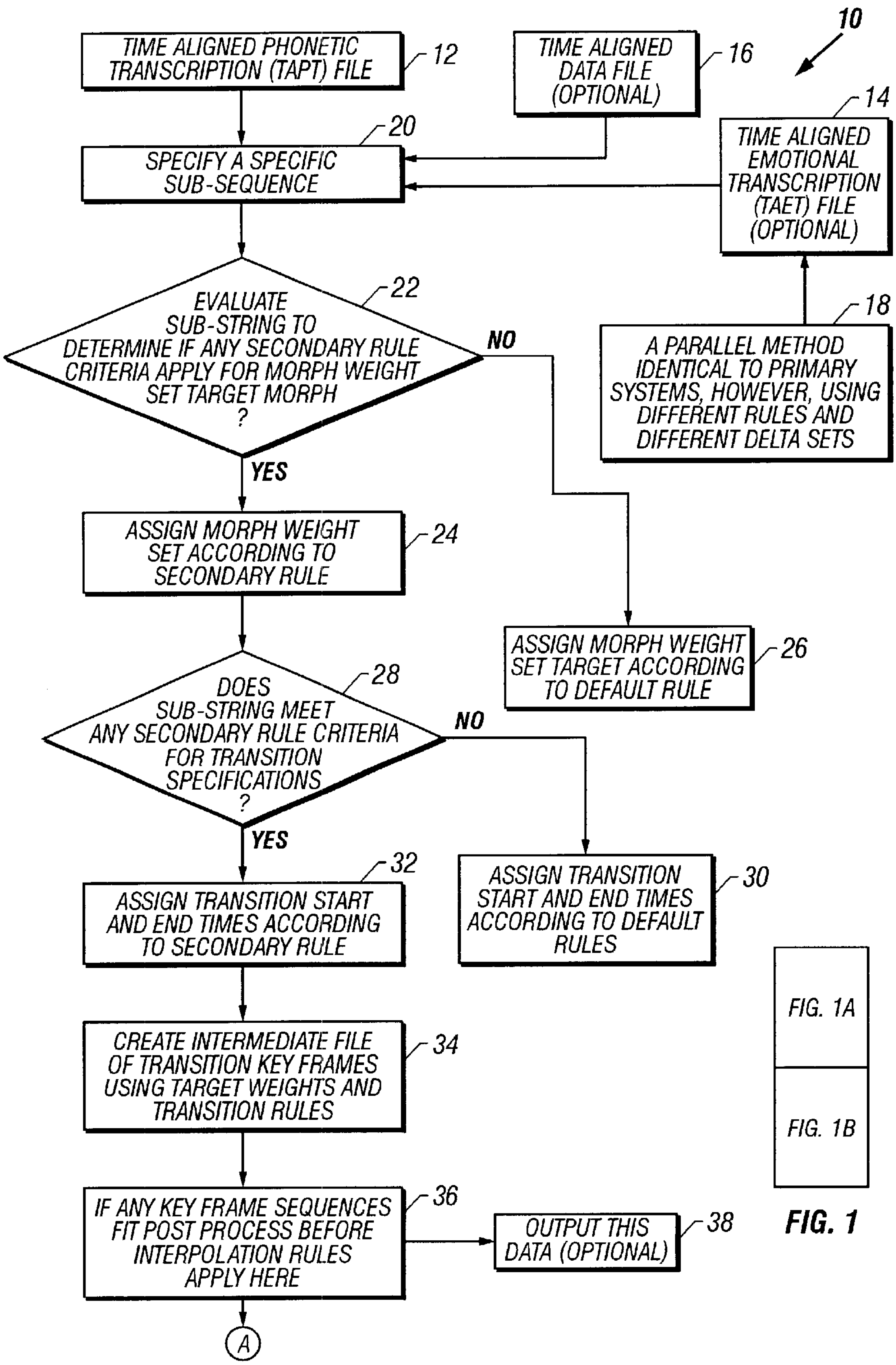

10
TIME ALIGNED PHONETIC TRANSCRIPTION (TAPT) FILE
12
TIME ALIGNED DATA FILE (OPTIONAL)
16
14
TIME ALIGNED EMOTIONAL TRANSCRIPTION (TAET) FILE (OPTIONAL)
20
SPECIFY A SPECIFIC SUB-SEQUENCE
22
EVALUATE SUB-STRING TO DETERMINE IF ANY SECONDARY RULE CRITERIA APPLY FOR MORPH WEIGHT SET TARGET MORPH ?
NO
YES
18
A PARALLEL METHOD IDENTICAL TO PRIMARY SYSTEMS, HOWEVER, USING DIFFERENT RULES AND DIFFERENT DELTA SETS
ASSIGN MORPH WEIGHT SET ACCORDING TO SECONDARY RULE
24
ASSIGN MORPH WEIGHT SET TARGET ACCORDING TO DEFAULT RULE
26
28
DOES SUB-STRING MEET ANY SECONDARY RULE CRITERIA FOR TRANSITION SPECIFICATIONS ?
NO
YES
32
ASSIGN TRANSITION START AND END TIMES ACCORDING TO SECONDARY RULE
ASSIGN TRANSITION START AND END TIMES ACCORDING TO DEFAULT RULES
30
CREATE INTERMEDIATE FILE OF TRANSITION KEY FRAMES USING TARGET WEIGHTS AND TRANSITION RULES
34
IF ANY KEY FRAME SEQUENCES FIT POST PROCESS BEFORE INTERPOLATION RULES APPLY HERE
36
OUTPUT THIS DATA (OPTIONAL)
38
A
FIG. 1A
FIG. 1B
FIG. 1

FIG. 1A

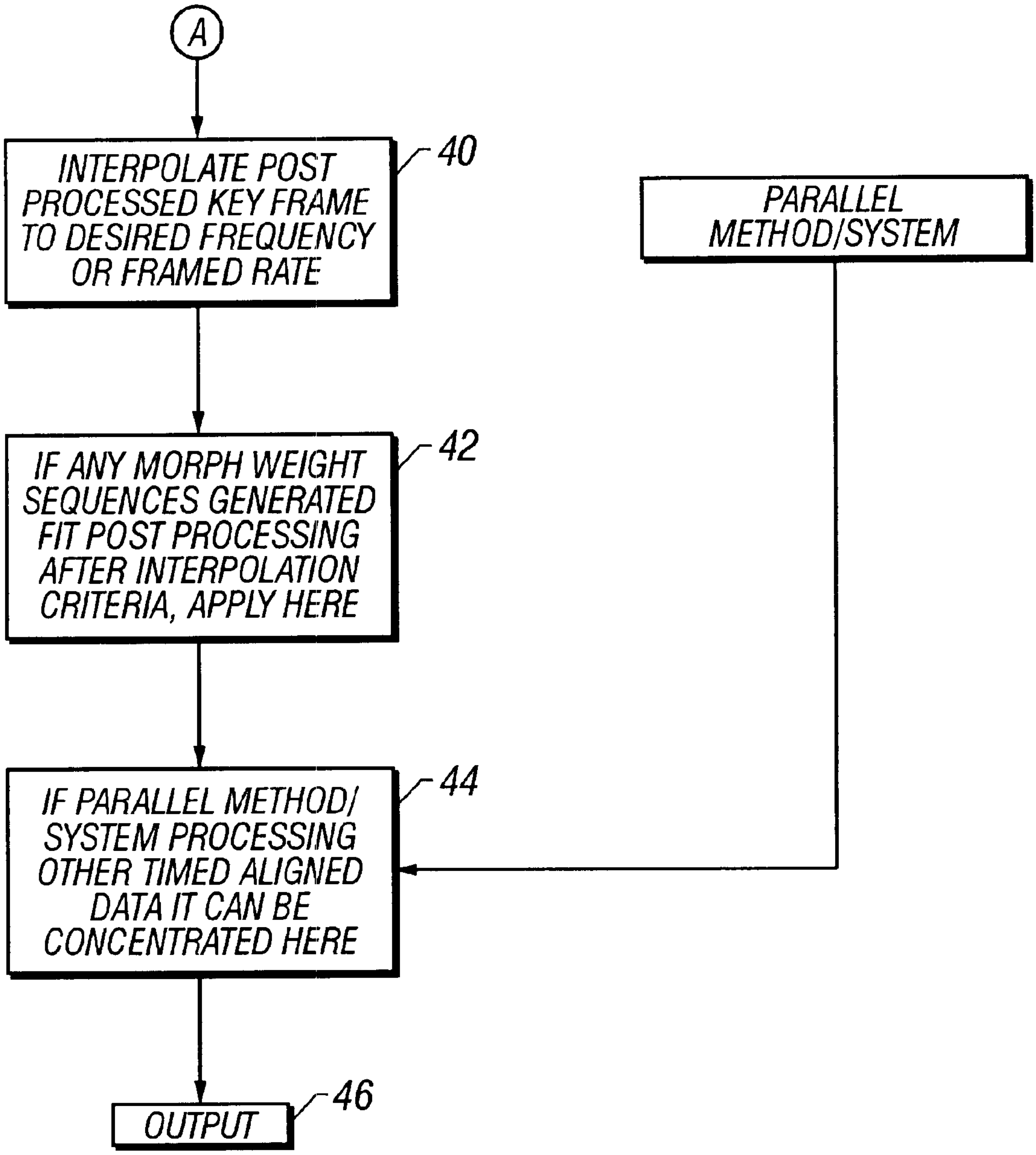

A
INTERPOLATE POST PROCESSED KEY FRAME TO DESIRED FREQUENCY OR FRAMED RATE
40
PARALLEL METHOD/SYSTEM
IF ANY MORPH WEIGHT SEQUENCES GENERATED FIT POST PROCESSING AFTER INTERPOLATION CRITERIA, APPLY HERE
42
IF PARALLEL METHOD/ SYSTEM PROCESSING OTHER TIMED ALIGNED DATA IT CAN BE CONCENTRATED HERE
44
OUTPUT
46

FIG. 1B

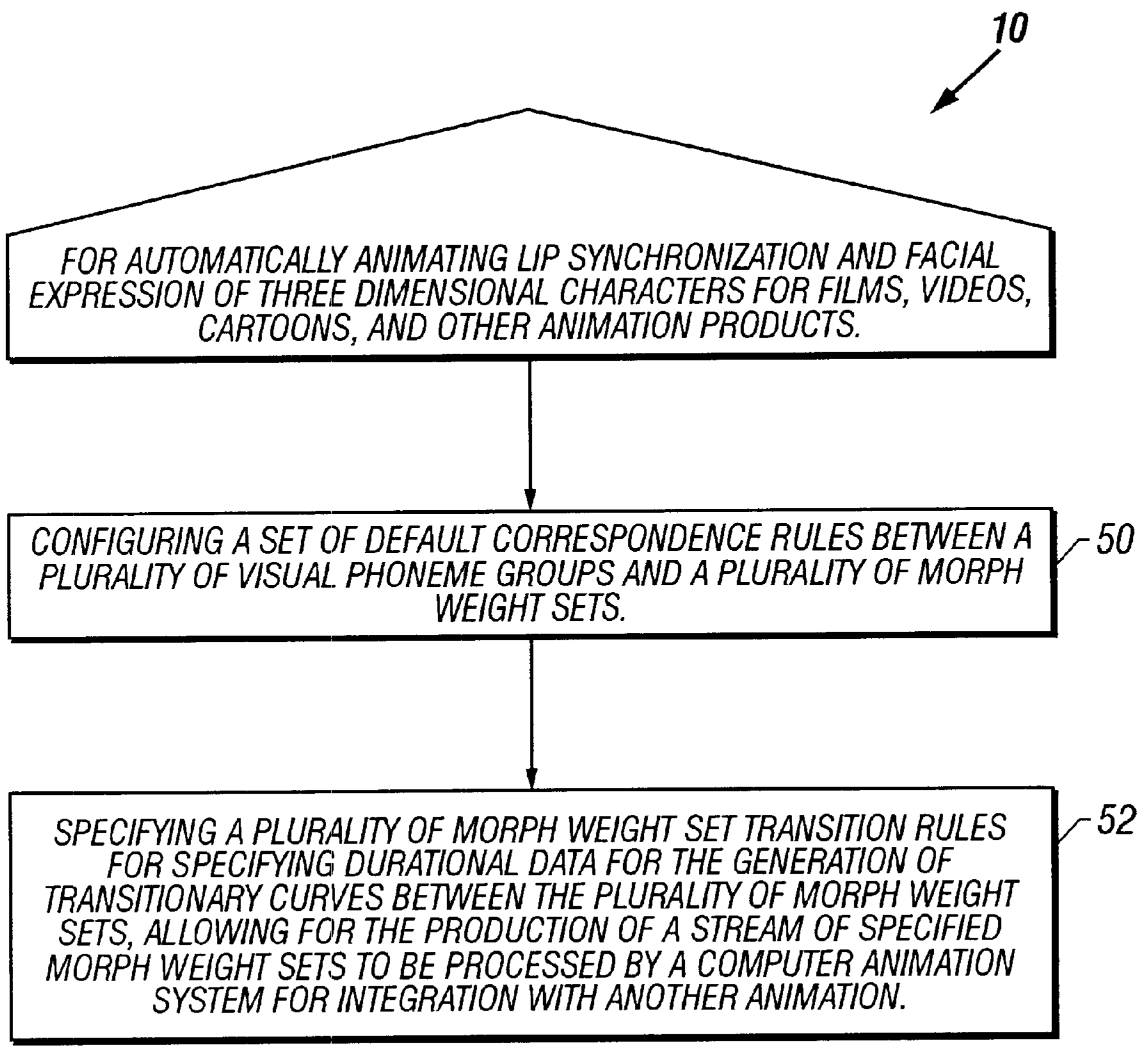

10
FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF THREE DIMENSIONAL CHARACTERS FOR FILMS, VIDEOS, CARTOONS, AND OTHER ANIMATION PRODUCTS.
CONFIGURING A SET OF DEFAULT CORRESPONDENCE RULES BETWEEN A PLURALITY OF VISUAL PHONEME GROUPS AND A PLURALITY OF MORPH WEIGHT SETS.
50
SPECIFYING A PLURALITY OF MORPH WEIGHT SET TRANSITION RULES FOR SPECIFYING DURATIONAL DATA FOR THE GENERATION OF TRANSITIONARY CURVES BETWEEN THE PLURALITY OF MORPH WEIGHT SETS, ALLOWING FOR THE PRODUCTION OF A STREAM OF SPECIFIED MORPH WEIGHT SETS TO BE PROCESSED BY A COMPUTER ANIMATION SYSTEM FOR INTEGRATION WITH ANOTHER ANIMATION.
52

FIG. 2

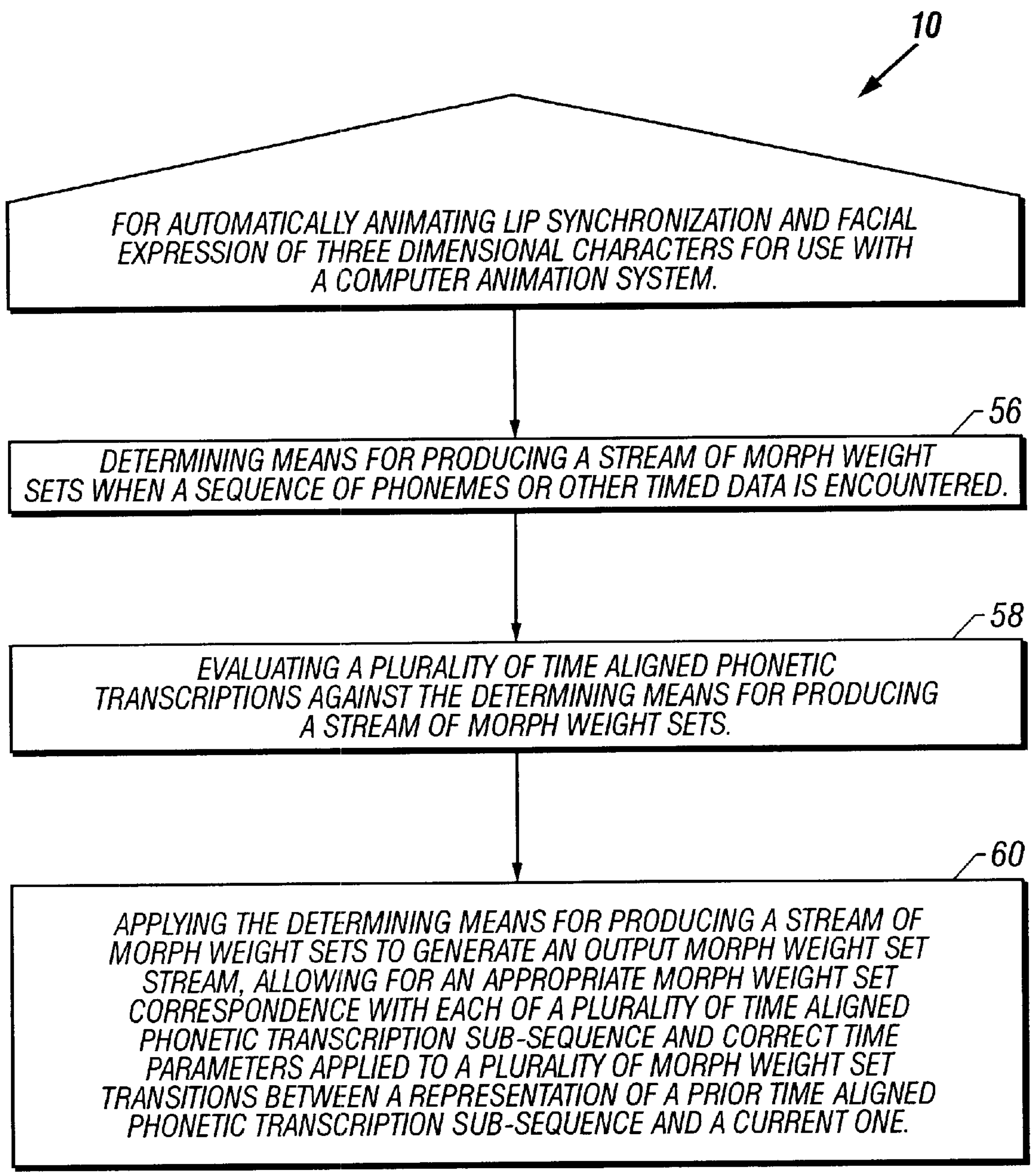

10
FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF THREE DIMENSIONAL CHARACTERS FOR USE WITH A COMPUTER ANIMATION SYSTEM.
56
DETERMINING MEANS FOR PRODUCING A STREAM OF MORPH WEIGHT SETS WHEN A SEQUENCE OF PHONEMES OR OTHER TIMED DATA IS ENCOUNTERED.
58
EVALUATING A PLURALITY OF TIME ALIGNED PHONETIC TRANSCRIPTIONS AGAINST THE DETERMINING MEANS FOR PRODUCING A STREAM OF MORPH WEIGHT SETS.
60
APPLYING THE DETERMINING MEANS FOR PRODUCING A STREAM OF MORPH WEIGHT SETS TO GENERATE AN OUTPUT MORPH WEIGHT SET STREAM, ALLOWING FOR AN APPROPRIATE MORPH WEIGHT SET CORRESPONDENCE WITH EACH OF A PLURALITY OF TIME ALIGNED PHONETIC TRANSCRIPTION SUB-SEQUENCE AND CORRECT TIME PARAMETERS APPLIED TO A PLURALITY OF MORPH WEIGHT SET TRANSITIONS BETWEEN A REPRESENTATION OF A PRIOR TIME ALIGNED PHONETIC TRANSCRIPTION SUB-SEQUENCE AND A CURRENT ONE.

FIG. 3

METHOD FOR AUTOMATICALLY ANIMATING LIP SYNCHRONIZATION AND FACIAL EXPRESSION OF ANIMATED CHARACTERS

This is a continuation of application Ser. No. 08/942,987 filed Oct. 2, 1997, now U.S. Pat. No. 6,307,576.

BACKGROUND OF THE INVENTION

1. Field of Invention

This invention relates generally to animation producing methods and apparatuses, and more particularly is directed to a method for automatically animating lip synchronization and facial expression for three dimensional characters.

2. Description of the Related Art

Various methods have been proposed for animating lip synchronization and facial expressions of animated characters in animated products such as movies, videos, cartoons, CD's, and the like. Prior methods in this area have long suffered from the need of providing an economical means of animating lip synchronization and character expression in the production of animated products due to the extremely laborious and lengthy protocols of such prior traditional and computer animation techniques. These shortcomings have significantly limited all prior lip synchronization and facial expression methods and apparatuses used for the production of animated products. Indeed, the limitations of cost, time required to produce an adequate lip synchronization or facial expression in an animated product, and the inherent imitations of prior methods and apparatuses to satisfactorily provide lip synchronization or express character feelings and emotion, leave a significant gap in the potential of animated methods and apparatuses in the current state of the art.

Time aligned phonetic transcriptions (TAPTS) are a phonetic transcription of a recorded text or soundtrack, where the occurrence in time of each phoneme is also recorded. A "phoneme" is defined as the smallest unit of speech, and corresponds to a single sound. There are several standard phonetic "alphabets" such as the International Phonetic Alphabet, and TIMIT created by Texas instruments, Inc. and MIT. Such transcriptions can be created by hand, as they currently are in the traditional animation industry and are called "x" sheets, or "gray sheets" in the trade. Alternatively such transcriptions can be created by automatic speech recognition programs, or the like.

The current practice for three dimensional computer generated speech animation is by manual techniques commonly using a "morph target" approach. In this practice a reference model of a neutral mouth position, and several other mouth positions, each corresponding to a different phoneme or set of phonemes is used. These models are called "morph targets". Each morph target has the same topology as the neutral model, the same number of vertices, and each vertex on each model logically corresponds to a vertex on each other model, or example, vertex #n on all models represents the left corner of the mouth, and although this is the typical case, such rigid correspondence may not be necessary.

The deltas of each vertex on each morph target relative to the neutral are computed as a vector from each vertex n on the reference to each vertex n on each morph target. These are called the delta sets. There is one delta set for each morph target.

In producing animation products, a value usually from 0 to 1 is assigned to each delta set by the animator and the value is called the "morph weight". From these morph

weights, the neutral's geometry is modified as follows: Each vertex N on the neutral has the corresponding delta set's vertex multiplied by the scalar morph weight added to it. This is repeated for each morph target, and the result summed. For each vertex v in the neutral model:

$$|\text{result}|=|\text{neutral}|+\Sigma_{x=1}{}^{n}|\text{delta setx}|\text{morph weigh}$$

where the symbol |xxx| is used to indicate the corresponding vector in each referenced set. For example, |result| is the corresponding resultant vertex to vertex v in the neutral model |neutral| and |delta setx| is the corresponding vector for delta set x.

If the morph weight of the delta set corresponding to the morph target of the character saying, for example, the "oh" sound is set to 1, and all others are set to 0, the neutral would be modified to look like the "oh target. If the situation was the same, except that the "oh" morph weight was 0.5, the neutral's geometry is modified half way between neutral and the "oh" morph target.

Similarly, if the situation was as described above, except "oh" weight was 0.3 and the "ee" morph weight was at 0.7, the neutral geometry is modified to have some of the "oh" model characteristics and more of the "ee" model characteristics. There also are prior blending methods including averaging the delta sets according to their weights.

Accordingly, to animate speech, the artist needs to set all of these weights at each frame to an appropriate value. Usually this is assisted by using a "keyframe" approach, where the artist sets the appropriate weights at certain important times ("keyframes") and a program interpolates each of the channels at each frame. Such keyframe approach is very tedious and time consuming, as well as inaccurate due to the large number of keyframes necessary to depict speech.

The present invention overcomes many of the deficiencies of the prior art and obtains its objectives by providing an integrated method embodied in computer software for use with a computer for the rapid, efficient lip synchronization and manipulation of character facial expressions, thereby allowing for rapid, creative, and expressive animation products to be produced in a very cost effective manner.

Accordingly, it is the primary object of this invention to provide a method for automatically animating lip synchronization and facial expression of three dimensional characters, which is integrated with computer means for producing accurate and realistic lip synchronization and facial expressions in animated characters. The method of the present invention further provides an extremely rapid and cost effective means to automatically create lip synchronization and facial expression in three dimensional animated characters.

Additional objects and advantages of the invention will be set forth in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The objects and advantages of the invention may be realized and obtained by means of the instrumentalities and combinations particularly pointed out in the appended claims.

SUMMARY OF THE INVENTION

To achieve the foregoing objects, and in accordance with the purpose of the invention as embodied and broadly described herein, a method is provided for controlling and automatically animating lip synchronization and facial expressions of three dimensional animated characters using weighted morph targets and time aligned phonetic transcrip-

tions of recorded text, and other time aligned data. The method utilizes a set of rules that determine the systems output comprising a stream or streams of morph weight sets when a sequence of timed phonemes or other timed data is encountered. Other timed data, such as pitch, amplitued, noise amounts, or emotional state data or emotemes such as "surprise, "disgust, "embarrassment", "timid smile", or the like, may be inputted to affect the output stream of morph weight sets.

The methodology herein described allows for automatically animating lip synchronization and facial expression of three dimensional characters in the creation of a wide variety of animation products, including but not limited to movies, videos, cartoons, CD's, software, and the like. The method and apparatuses herein described are operably integrated with computer software and hardware.

In accordance with the present invention there also is provided a method for automatically animating lip synchronization and facial expression of three dimensional characters for films, videos, cartoons, and other animation products, comprising configuring a set of default correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and specifying a plurality of morph weight set transition rules for specifying durational data for the generation of transitionary curves between the plurality of morph weight sets, allowing for the production of a stream of specified morph weigh sets to be processed by a computer animation system for integration with other animation, whereby animated lip synchronization and facial expression of animated characters may be automatically controlled and produced.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate a preferred embodiment of the invention and, together with a general description given above and the detailed description of the preferred embodiment given below, serve to explain the principles of the invention.

FIG. **1** is a flow chart showing the method of the invention with an optional time aligned emotional transcription file, and another parallel timed data file, according to the invention.

FIG. **2** is a flow chart illustrating the principal steps of the present method, according to the invention.

FIG. **3** is another representational flow chart illustrating the present method, according to the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Reference will now be made in detail to the present preferred embodiments of the invention as illustrated in the accompanying drawings.

In accordance with the present invention, there is provided as illustrated in FIGS. **1–3**, a method for controlling and automatically animating lip synchronization and facial expressions of three dimensional animated characters using weighted morph targets and time aligned phonetic transcriptions of recorded text. The method utilizes a set of rules that determine the systems output comprising a stream of morph weight sets when a sequence of timed phonemes is encountered. Other timed data, such as timed emotional state data or emotemes such as "surprise, "disgust, "embarrassment", "timid smile", pitch, amplitued, noise amounts or the like, may be inputted to affect the output stream of morph weight sets.

The method comprises, in one embodiment, configuring a set of default correspondence rules between a plurality of visual phoneme groups and a plurality of morph weight sets; and specifying a plurality of morph weight set transition rules for specifying durational data for the generation of transitionary curses between the plurality of morph weight sets, allowing for the production of a stream of specified morph weight sets to be processed by a computer animation system for integration with other animation, whereby animated lip synchronization and facial expression of animated characters may be automatically produced.

There is also provided, according to the invention a method for automatically animating lip synchronization and facial expression of three dimensional characters for use with a computer animation system, comprising the steps of: determining means for producing a stream of morph weight sets when a sequence of phonemes is encountered; evaluating a plurality of time aligned phonetic transcriptions or other timed data such as pitch, amplitude, noise amounts and the like, against the determining means for producing a stream of morph weight sets; applying said determining means for producing a stream of morph weight sets to generate an output morph weight set stream, allowing for an appropriate morph weight set correspondence with each of a plurality of time aligned phonetic transcription sub-sequences and correct time parameters applied to a plurality of morph weight set transitions between a representation of a prior time aligned phonetic transcription subsequence and a current one, whereby lip synchronization and facial expressions of animated characters is automatically controlled and produced.

The method preferably comprises a set of rules that determine what the output morph weight set stream will be when any sequence or phonemes and their associated times is encountered. As used herein, a "morph weight set" is a set of values, one for each delta set, that, when applied as described, transform the neutral mode to some desired state, such as speaking the "oo" sound or the "th" sound. Preferably, one model id designated as the anchor model, which the deltas are computed in reference to. If for example, the is a morph target that represents all possible occurrences of an "e" sound perfectly, it's morph weight set would be all zeros for all delta sets except for the delta set corresponding to the "ee" morph target, which would set to 1.

Preferably, each rule comprises two parts, the rule's criteria and the rule's function. Each sub-sequence of time aligned phonetic transcription (TAPT) or other timed data such as pitch, amplitude, noise amount or the like, is checked against a rule's criteria to see if that rule is applicable. If so, the rule's function is applied to generate the output. The primary function of the rules is to determined 1) the appropriate morph weight set correspondence with each TAPT sub-sequence; and 2) the time parameters of the morph weight set transitions between the representation of the prior TAPT sub-sequence or other timed data and the current one. Conditions 1) and 2) must be completely specified for any sequence of phonemes and times encountered. Together, such rules are used to create a continuous stream of morph weight sets.

In the present method, it is allowable for more than one phoneme to be represented by the same morph target, for example, "sss" and "zzz". Visually, these phonemes appear similar. Through the use of such rules, the user can group phonemes together that have a similar visual appearance into visual phonemes" that function the same as one another. It is also acceptable, through the rules, to ignore certain

phoneme sequences. For example, a rule could specify: "If in the TAPT, there are two or more adjacent phonemes that are in the same "visual phoneme" group, all but the first are ignored".

The rules of the present method may be categorized in three main groupings; default rules, auxiliary rules and post processing rules. The default rules must be complete enough to create valid output for any TAPT encountered at any point in the TAPT. The secondary rules are used in special cases; for example, to substitute alternative morph weight set correspondences and/or transition rules if the identified criteria are met. The post processing rules are used to further manipulate the morph weight set stream after the default or secondary rules are applied, and can further modify the members of the morph weight sets determined by the default and secondary rules and interpolation.

If for example, specific TAPT subsequence does not fit the criteria for any secondary rules, then the default rules take effect. If, on the other hand, the TAPT sub-sequence does fit the criteria for a secondary rule(s) they take precedence over the default rules. A TAPT sub-sequence take into account the current phoneme and duration, and a number of the preceding and following phonemes and duration's as well may be specified.

Preferably, the secondary rules effect morph target correspondence and weights, or transition times, or both. Secondary rules can create transitions and correspondences even where no phoneme transitions exist. The secondary rules can use as their criteria the phoneme, the duration or the phoneme's context in the output stream, that is what phonemes are adjacent or in the neighborhood to the current phoneme, what the adjacent duration's are, and the like.

The post processing rules are preferably applied after a preliminary output morph weight set is calculated so as to modify it. Post processing rules can be applied before interpolation and/or after interpolation, as described later in this document. Both the secondary and post processing rules are optional, however, they may in certain applications be very complex, and in particular circumstances contribute more to the output than the default rules.

In FIG. **1**, a flow chart illustrates the preferred steps of the methodology **10** or automatically animating lip synchronization and facial expression of three dimensional animated characters of the present invention. A specific sub-sequence **20** is selected from the TAPT file **12** and is evaluated **22** to determine if any secondary rule criteria for morph weight set target apply. Time aligned emotional transcription file **14** data may be inputted or data from an optional time aligned data file **16** may be used. Also shown is a parallel method **18** which may be configured identical to the primary method described, however, using different timed data rules and different delta sets. Sub-sequence **20** is evaluated **22** to determine if any secondary rule criteria apply. If yes, then a morph weight set is assigned **24** according to the secondary rules, if no, then a morph weight set is assigned **26** according to the default rules. If the sub-string meets any secondary rule criteria for transition specification **28** then a transition start and end time are assigned according to the secondary rules **32**, if no, then assign transition start and end times **30** according to default rules. Then an intermediate file of transition keyframes using target weights and transition rules as generated are created **34**, and if any keyframe sequences fit post process before interpolation rules they are applied here **36**. This data may be output **38** here if desired. If not, then interpolate using any method post processed keyframes to a desired frequency or frame rate **40** and if any morph weight sequences generated fit post processing after interpolation criteria, they are applied **42** at this point. If parallel methods or systems are used to process other timed aligned data, they may be concatenated here **44**, and the data output **46**.

In FIG. **2**, the method for automatically animating lip synchronization and facial expression of three dimensional characters for films, videos, cartoons, and other animation products **10** is shown according to the invention, where box **50** show the step of configuring a set of default correspondence rules between a plurality of visual phoneme groups or other timed input data and a plurality of morph weight sets. Box **52** shows the steps of specifying a plurality of morph weight set transition rules for specifying durational data for the generation of transitionary curves between the plurality of morph weight sets, allowing for the production of a stream of specified morph weight sets to be processed by a computer animation system for integration with other animation, whereby animated lip synchronization and facial expression of animated characters may be automatically produced.

With reference now to FIG. **3**, method **10** for automatically animating lip synchronization and facial expression of three dimensional characters for use with a computer animation system is shown including box **56** showing the step of determining means for producing a stream of morph weight sets when a sequence of phonemes is encountered. Box **53**, showing the step of evaluating a plurality of time aligned phonetic transcriptions or other timed ata such as pitch, amplitude, noise amounts, and the like, against said determining means for producing a stream of morph weight sets. In box **60** the steps of applying said determining means for producing a stream of morph weight sets to generate an output morph weight set stream, allowing for an appropriate morph weight set correspondence with each of a plurality of time aligned phonetic transcription sub-sequences and correct time parameters applied to a plurality of morph weight set transitions between a representation of a prior time aligned phonetic transcription sub-sequence and a current one, whereby lip synchronization and facial expressions of animated characters is automatically controlled and produced are shown according to the invention.

In operation and use, the user must manually set up default correspondence rules between all visual phoneme groups and morph weight sets. To do this, the user preferably specifies the morph weight sets which correspond to the model speaking, for example the "oo" sound, the "th" sound, and the like. Next, default rules must be specified. These rules specify the durational information needed to generate appropriate transitionary curves between morph weight sets, such as transition start and end times. "transition" between two morph weigh sets is defined as each member of the morph weight set transitions from it's current state to it's target state, starting at the transition start time and ending at the transition end time. The target state is the morph weight set determined by a correspondence rule.

The default correspondence rules and the default morph weight set transition rules define the default system behavior. If all possible visual phoneme groups or all members of alternative data domains have morph weight set correspondence, any phoneme sequence can be handled with this rule set alone. However, additional rules are desirable for effects, exceptions, and uniqueness of character, as further described below.

According to the method of the invention, other rules involving phoneme's duration and/or context can be speci-

fied. Also, any other rules that do not fit easily into the above mentioned categories can be specified. Examples of such rules are described in greater detail below and are termed the "secondary rules". If a timed phoneme or sub-sequence of timed phonemes do not fit the criteria for any of the secondary rules, the default rules are applied as seen in FIG. **1**.

It is seen that through the use or these rules, an appropriate morph weight stream is produced. The uninterpolated morph weight stream has entries only at transition start and end time, however. These act as keyframes. A morph weight set may be evaluated at any time by interpolating between these keyframes, using conventional methods. This is how the output stream is calculated each desired time frame. For example, for television productions, the necessary resolution is 30 evaluations per second.

The post processing rules may be applied either before or after the above described interpolation step, or both. Some rules may apply only to keyframes before interpolation, some to interpolated data. If applied before the interpolation step, this affects the keyframes. if applied after, it effects the interpolated data. Post processing can use the morph weight sets calculated by the default and secondary rules. Post processing rules can use the morph weigh sets or sequences as in box **44** of FIG. **1**, calculated by the default and secondary rules. Post processing rules can modify the individual members of the morph weight sets previously generated. Post processing rules may be applied in addition to other rules, including other post processing rules. Once the rule set up is completed as described, the method of the present invention can take any number and length TAPT's as input, and automatically output the corresponding morph weight set stream as seen in FIGS. **1–3**.

For example, a modeled neutral geometric representation of a character for an animated production such as a movies, video, cartoon, CD or the like, with six morph targets, and their delta sets determined. Their representations, for example, are as follows:

Delta Set	Visual Representation
1	"h"
2	"eh"
3	"1"
4	"oh"
5	exaggerated "oh"
6	special case "eh" used during a "snide laugh" sequences

In this example, the neutral model is used to represent silence. The following is an example of a set of rules, according to the present method, of course this is only an example of a set of rules which could be use for illustrative purposes, and many other rules could be specified according to the method of the invention.

Default Rules:

Default Correspondence Rules

Criteria: Encounter a "h" as in "house"

Function: Use morph weight set (1,0,0,0,0,0) as transition target.

Criteria: Encounter an "eh" as in "bet"

Function: Use morph weight set (0,1,0,0,0,0) as transition target.

Criteria: Encounter a "l" as in "old"

Function: Use morph weight set (0,0,1,0,0,0) as transition target.

Criteria: Encounter an "oh" as in "old"

Function: Use morph weight set (0,0,0,1,0,0) as transition target.

Criteria: encounter a "silence"

Function: use morph weight set (0,0,0,0,0,0) as transition target.

Default Transition Rule

Criteria: Encounter any phoneme

Function: Transition start time=(the outgoing phoneme's end time)−0.1*(the outgoing phoneme's duration); transition end time=(the incoming phoneme's start time)+ 0.1* (the incoming phoneme's duration)

Secondary Rules

Criteria: Encounter an "oh" with a duration greater than 1.2 seconds.

Function: Use morph weigh set (0,0,0,0,1,0)

Criteria: Encounter and "eh" followed by an "h" and preceded by an "h".

Function: Use morph weigh set (0,0,0,0,0,1) as transition target.

Criteria: Encounter any phoneme preceded by silence

Function: Transition start time=(the silence's end time)−0.1*(the incoming phoneme's duration) Transition end time=the incoming phoneme's start time

Criteria: Encounter silence preceded by any phoneme.

Function: Transition start time=the silence's start time−0.1* (the outgoing phoneme's duration)

Post Processing Rules

Criteria: Encounter a phoneme duration under 0.22 seconds.

Function: Scale the transition target determined by the default and secondary rules by 0.8 before interpolation.

Accordingly, using this example, if the user were to use these rules for the spoken word "Hello", at least four morph targets and a neutral target would be required, that is, one each for the sound of "h", "e", "l","oh " and their associated delta sets. For example, a TAPT representing the spoken word "hello" could be configured as,

Time	Phoneme
0.0	silence begins
0.8	silence ends, "h" begins
1.0	"h" ends, "eh" begins
1.37	"eh" ends, "1" begins
1.6	"1" ends, "oh" begins
2.1	"oh" ends, silence begins.

The method, for example embodied in computer software for operation with a computer or computer animation system would create an output morph weight set stream as follows:

Time	D.S.1("h")	D.S.2("eh")	D.S.3("l")	D.S.4("oh")	D.S.5(aux"oh")	D.S.6
0.0	0	0	0	0	0	0
0.78	0	0	0	0	0	0

-continued

Time	D.S.1("h")	D.S.2("eh")	D.S.3("l")	D.S.4("oh")	D.S.5(aux"oh")	D.S.6
0.8	1	0	0	0	0	0
0.98	1	0	0	0	0	0
1.037	0	1	0	0	0	0
1.333	0	1	0	0	0	0
1.403	0	0	1	0	0	0
1.667	0	0	1	0	0	0
1.74	0	0	0	1	0	0
2.1	0	0	0	1	0	0
2.14	0	0	0	0	0	0

Such morph weight sets act as keyframes, marking the transitionary points. A morph weight set can be obtained for any time within the duration of the TAPT by interpolating between the morph weight sets using conventional methods well known in the art. Accordingly, a morph weight set can be evaluated at every frame. However, the post processing rules can be applied to the keyframes before interpolation as in box **36** of FIG. **1**, or to the interpolated data as in box **40** of FIG. **1**. From such stream of morph weight sets, the neutral model is deformed as described above, and then sent to a conventional computer animation system for integration with other animation. Alternatively, the morph weight set stream can be used directly by an animation program or package, wither interpolated or not.

The rules of the present invention are extensible and freeform in the sense that they may be created as desired and adapted to a wide variety of animation characters, situations, and products. As each rule comprise a criteria and function, as in an "if . . . then . . . else" construct. The following are illustrative examples of other rules which may be used with the present methodology.

For example, use {0,0,0,0 . . . } as the morph weighs, set when a "m" is encountered. This is a type of default rule, where: Criteria: Encounter a "m" phoneme of any duration. Function: Use a morph weight set {0,0,0,0 . . . 0} as a transition target.

Another example would be creating several slightly different morph targets for each phoneme group, and using them randomly each time that phoneme is spoken. This would give a more random, or possibly comical or interesting look to the animation's. This is a secondary rule.

An example of post processing rule, before interpolation would be to add a small amount of random noise to all morph weight channels are all keyframes. This would slightly alter the look of each phoneme to create a more natural look.

Criteria: Encounter any keyframe

Function: Add a small random value to each member of the morph weight set prior to interpolation.

An example of a post processing rule, after interpolation would be to add a component of an auxiliary morph target (one which does not correspond directly to a phoneme) to the output stream in a cyclical manner over time, after interpolation. If the auxiliary morph target had the character's mouth moved to the left, for example, the output animation would have the character's mouth cycling between center to left as he spoke.

Criteria: Encounter any morph weight set generated by interpolation

Function: Add a value calculated through a mathematical expression to the morph weigh set's member that corresponds to the auxiliary morph target's delta set weight. The expression might be, for example: 0.2*sin (0.2*time*2*pi)+0.2. This rule would result in an oscillation of the animated character's mouth every five seconds.

Another example of a secondary rule is to use alternative weight sets (or morph weight set sequences) for certain contexts of phonemes, for example, in an "oh" is both preceded and followed by an "ee" then use an alternate "oh". This type of rule can make speech idiosyncrasies, as well as special sequences for specific words (which are a combination of certain phonemes in a certain context). This type of rule can take into consideration the differences in mouth positions for similar phonemes based on context. For example, the "l" in "hello" is shaped more widely than the "l" in "burly" due to it's proximity to an "eh" as opposed to a "r".

Criteria: Encounter an "l" preceded by an "r".

Function: Use a specified morph weight set as transition target.

Another secondary rule could be, by way of illustration, that if a phoneme is longer than a certain duration, substitute a different morph target. this can add expressiveness to extended vowel sounds, for instance, if a character says "HELLOOOOOOO!" a more exaggerated "oh" model would be used.

Criteria: Encounter an "oh" longer than 0.5 seconds and less than 1 second.

Function: Use a specified morph weight set as a transition target.

If a phoneme is longer than another phoneme of even longer duration, a secondary rule may be applied to create new transitions between alternate morph targets at certain intervals, which may be randomized, during the phoneme's duration. This will add some animation to extremely long held sounds, avoiding a rigid look. This is another example of a secondary rule

Criteria: Encounter an "oh" longer than 1 second long.

Function: Insert transitions between a defined group of morph weight sets at 0.5 second intervals, with transition duration's of 0.2 seconds until the next "normal" transition start time is encountered.

If a phoneme is shorter than a certain duration, its corresponding morph weight may be scaled by a factor smaller than 1. This would create very short phonemes not appear over articulated. Such a post processing rule, applied before interpolation would comprise:

Criteria: Encounter a phoneme duration shorter than 0.1 seconds.

Function: Multiply all members of the transition target (already determined by default and secondary rules by duration/0.1.

As is readily apparent a wide variety of other rules can be created to add individuality to the different characters.

A further extension of the present method is to make a parallel method or system, as depicted in box **14** of FIG. **1**, that uses time aligned emotional transcriptions (TAET) that correspond to facial models of those emotions. Using the same techniques as previously described additional morph

weight set streams can be created that control other aspects of the character that reflect facial display of emotional state. Such morph weight set streams can be concatenated with the lip synchronization stream. In addition, the TAET data can be used in conjunction with the lip synchronization secondary rules to alter the lip synchronization output stream. For example:

Criteria: An "L" is encountered in the TAPT and the nearest "emoteme" in the TAET is a "smile".

Function: Use a specified morph weight set as transition target.

As is evident from the above description, the automatic animation lip synchronization and facial expression method described may be used on a wide variety of animation products. The method described herein provides an extremely rapid, efficient, and cost effective means to provide automatic lip synchronization and facial expression in three dimensional animated characters. The method described herein provides, for the first time, a rapid, effective, expressive, and inexpensive means to automatically create animated lip synchronization and facial expression in animated characters. The method described herein can create the necessary morph weight set streams to create speech animation when given a time aligned phonetic transcription of spoken text and a set of user defined rules for determining appropriate morph weight sets for a given TAPT sequence. This method also defines rules describing a method of transitioning between these sets through time. The present method is extensible by adding new rules, and other timed data may be supplied, such as time "emotemes" that will effect the output data according to additional rules that take this data into account. In this manner, several parallel systems may be used on different types of timed data and the results concatenated, or used independently. Accordingly, additional advantages and modification will readily occur to those skilled in the art. The invention in its broader aspects is, therefore, not limited to the specific methodological details, representative apparatus and illustrative examples shown and described. Accordingly, departures from such details may be made without departing from the spirit or scope of the applicant's inventive concept.

What is claimed is:

1. A method for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:

obtaining a first set of rules that defines a morph weight set stream as a function of phoneme sequence and times associated with said phoneme sequence;

obtaining a plurality of sub-sequences of timed phonemes corresponding to a desired audio sequence for said three-dimensional characters;

generating an output morph weight set stream by applying said first set of rules to each sub-sequence of said plurality of sub-sequences of timed phonemes; and

applying said output morph weight set stream to an input sequence of animated characters to generate an output sequence of animated characters with lip and facial expression synchronized to said audio sequence.

2. The method of claim **1**, wherein said first set of rules comprises:

correspondence rules between all visual phoneme groups and morph weight sets; and

morph weight set transition rules specifying durational data between morph weight sets.

3. The method of claim **2**, wherein said durational data comprises transition start and transition end times.

4. The method of claim **1**, wherein said desired audio sequence is from a pre-recorded live performance.

5. The method of claim **1**, wherein said desired audio sequence is synthetically generated by a computer.

6. The method of claim **1**, wherein said plurality of subsequences of timed phonemes is obtained from a file.

7. The method of claim **1**, wherein said plurality of subsequences of timed phonemes is generated during animation.

8. The method of claim **1**, wherein said output sequence of animated characters is transmitted over a computer network.

9. The method of claim **1**, wherein said generating said output morph weight stream comprises:

generating an appropriate morph weight set corresponding to each subsequence of said timed phonemes; and

generating time parameters for transition of said appropriate morph weight set from a morph weight set of a prior sub-sequence of said timed data.

10. The method of claim **1**, wherein each of said first set of rules comprises a rule's criteria and a rule's function.

11. The method of claim **10**, wherein said generating an output morph weight set stream comprises:

checking each sub-sequence of said plurality of sub-sequences of timed data for compliance with said rule's criteria; and

generating an output morph weight set and transition parameters by applying said rule's function upon said compliance with said criteria.

12. The method of claim **1**, wherein said first set of rules comprises a default set of rules and an optional secondary set of rules, said secondary set of rules having priority over said default set of rules.

13. The method of claim **1**, wherein said plurality of subsequences of timed phonemes comprises a timed aligned phonetic transcriptions sequence.

14. The method of claim **1**, wherein said plurality of subsequences of timed phonemes comprises time aligned data.

15. The method of claim **13**, wherein said plurality of subsequences of timed phonemes further comprises time aligned emotional transcription data.

16. The method of claim **9**, wherein said transition parameters comprises:

transition start time; and

transition end time.

17. The method of claim **16**, further comprising:

generating said output morph weight set stream by interpolating between morph weight sets at said transition start time and said transition end time according to a desired frame rate of said output sequence of animated characters.

18. The method of claim **1**, further comprising:

applying a second set of rules to said output morph weight set prior to said generating of said output sequence of animated characters.

19. An apparatus for automatically animating lip synchronization and facial expression of three-dimensional characters comprising:

a computer system;

computer code in said computer system, said computer code comprising:

a method for obtaining a first set of rules that defines a morph weight set stream as a function of phoneme sequence and times associated with said phoneme sequence;

a method for obtaining a plurality of sub-sequences of timed phonemes corresponding to a desired audio sequence for said three-dimensional characters;

a method for generating an output morph weight set stream by applying said first set of rules to each sub-sequence of said plurality of subsequences of timed phonemes;

a method for applying said output morph weight set stream to an input sequence of animated characters to generate an output sequence of animated characters with lip and facial expression synchronized to said audio sequence.

20. The apparatus of claim **19**, wherein said first set of rules comprises:

correspondence rules between all visual phoneme groups and morph weight sets; and

morph weight set transition rules specifying durational data between morph weight sets.

21. The apparatus of claim **20**, wherein said durational data comprises transition start and transition end times.

22. The apparatus of claim **19**, wherein said desired audio sequence is from a pre-recorded live performance.

23. The apparatus of claim **19**, wherein said desired audio sequence is synthetically generated by a computer.

24. The apparatus of claim **19**, said plurality of subsequences of timed phonemes is obtained from a file.

25. The apparatus of claim **19**, wherein said plurality of subsequences of timed phonemes is generated during animation.

26. The apparatus of claim **19**, wherein said output sequence of animated characters is transmitted over a computer network.

27. The apparatus of claim **19**, wherein said generating said output morph weight stream comprises:

generating an appropriate morph weight set corresponding to each subsequence of said timed phonemes; and

generating time parameters for transition of said appropriate morph weight set from a morph weight set of a prior sub-sequence of said timed data.

28. The apparatus of claim **19**, wherein each of said first set of rules comprises a rule's criteria and a rule's function.

29. The apparatus of claim **28**, wherein said generating an output morph weight set stream comprises:

checking each sub-sequence of said plurality of subsequences of timed data for compliance with said rule's criteria; and

generating an output morph weight set and transition parameters by applying said rule's function upon said compliance with said criteria.

30. The apparatus of claim **19**, wherein said first set of rules comprises a default set of rules and an optional secondary set of rules, said secondary set of rules having priority over said default set of rules.

31. The apparatus of claim **19**, wherein said plurality of subsequences of timed phonemes comprises a timed aligned phonetic transcriptions sequence.

32. The apparatus of claim **19**, wherein said plurality of subsequences of timed phonemes comprises time aligned data.

33. The apparatus of claim **31**, wherein said plurality of subsequences of timed phonemes further comprises time aligned emotional transcription data.

34. The apparatus of claim **27**, wherein said transition parameters comprises:

transition start time; and

transition end time.

35. The apparatus of claim **34**, wherein said computer code further comprises:

a method for generating said output morph weight set stream by interpolating between morph weight sets at said transition start time and said transition end time according to a desired frame rate of said output sequence of animated characters.

36. The apparatus of claim **19**, wherein said computer code further comprises:

a method for applying a second set of rules to said output morph weight set prior to said generating of said output sequence of animated characters.

* * * * *

CERTIFICATE OF SERVICE

I certify that today, March 3, 2015, I electronically filed the foregoing Corrected Brief for Plaintiff-Appellant McRO, Inc. with the Clerk of the Court for the U.S. Court of Appeals for the Federal Circuit using the appellate CM/ECF system. Counsel of record for all parties will be served by the appellate CM/ECF system.

March 3, 2015 /s/ Jeffrey A. Lamken

CERTIFICATE OF COMPLIANCE

1. This brief complies with the type-volume limitation of Fed. R. App. P. 32(a)(7)(B) because this brief contains 13,647 words, excluding the parts of the brief exempted by Fed. R. App. P. 32(a)(7)(B)(iii) and Fed. Cir. R. 32(b).

2. This brief complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type style requirements of Fed. R. App. P. 32(a)(6) because this brief has been prepared in a proportionally spaced typeface using Microsoft Word in Times New Roman 14 point font.

February 27, 2015 /s/ Jeffrey A. Lamken